Processing of Alternative DNA Structures in the Human Telomere

by

Gerald Joseph Nora

B.Phil.,B.A., University of Pittsburgh, 2002

Submitted to the Graduate Faculty of

School of Medicine in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy

University of Pittsburgh

2010

UNIVERSITY OF PITTSBURGH

School of Medicine

This dissertation was presented

by

Gerald J. Nora

It was defended on

March 4, 2010

and approved by

Saleem Khan, Professor, Biochemistry and Molecular Genetics

Sanford Leuba, Associate Professor, Cell Biology and Physiology

Michael Trakselis, Assistant Professor, Chemistry

Dissertation Advisor: Patricia L. Opresko, Assistant Professor, Environmental and Occupational Health

Processing of Alternative DNA Structures in the Human Telomere

Gerald J. Nora, PhD

University of Pittsburgh, 2010

Telomeres help maintain the overall genomic stability of an organism, and telomeric homeostasis is critical to navigating between aging and cancer. Telomeric dysfunction is implicated as a contributing factor in numerous aging-related diseases, such as diabetes, impaired hematopoeisis, and atherosclerosis. Telomeric homeostasis is maintained by a shelterin complex of six proteins and an array of telomere-associated proteins that interact with the central shelterin complex, such as the Werner syndrome helicase/exonuclease protein (WRN) or p53. Telomeres also have non-canonical DNA structures that are critical towards their function, especially G-quadruplex DNA (G4 DNA) and Holliday Junctions. The former are pseudoknots that form on the G-rich 3' single-stranded tail of the telomere and may block telomere replication and lengthening, when the 3' telomeric tail is exposed in the "open" conformation. We found that the shelterin protein protection of telomeres 1 (POT1) competes with and destabilizes G4 DNA on a physiologically realistic telomeric tail substrate, leading to an equilibrium population of diminished G4 DNA coexisting with POT1. While POT1 is a passive binder of DNA, the destabilizing effect of bound POT1 on pre-existing G4 DNA leads to an emergent, *de facto* cooperativity in G4 DNA unfolding by POT1.

Holliday Junctions (HJ) form when the telomere is in a "closed" conformation, in which the 3' telomeric tail invades the duplex telomeric DNA, creating a displacement loop (D-loop) and sequestering the end of the chromosome from unwanted DNA damage responses. The D-

loop is a homologous recombination intermediate, and we demonstrate that telomere repeat binding factor 2 (TRF2) is necessary to protect HJ DNA from unwanted WRN helicase activity, which has been thought to branch migrate the D-loop into a target for Holliday Junction cleaving enzymes, causing sudden telomere shortening. TRF2 also protects HJ DNA from Holliday Junction cleaving enzymes, and the cleavage protection is due largely to the HJ-binding B-domain on TRF2. In contrast, we found that TRF2-mediated protection against WRN depends on both the B domain and the telomeric-repeat binding Myb domain. We have therefore discovered an overlapping but distinct role for TRF2 in maintaining telomeric stability.

Our work has elucidated novel structural and functional data on the modulation of non-canonical DNA structures by shelterin and telomere-associated proteins. These data help us elucidate the mechanisms underlying cellular and animal models for telomere instability and aging.

TABLE OF CONTENTS

ABBREVIATIONS

AFM	Atomic Force Microscopy
ATM	Ataxia Telangiectasia Mutated
ATR	ATM and Rad3 related
BME	Beta-mercaptoethanol
BSA	Bovine Serum Albumin
CD	Circular Dichroism
D-loop	Displacement loop
dsDNA	Double stranded DNA
DTT	Dithiothreitol
GEN1	XPG-like endonuclease 1
G4 DNA	G-Quadruplex DNA
HJ	Holliday Junction
HR	Homologous Recombination
HRDC	Helicase and RNAse D C-terminal
NHEJ	Non-Homologous End Joining
NMR	Nuclear Magnetic Resonance
POT1	Protection of Telomeres 1
PDB	Protein Data Bank
RAP1	Transcriptional Repressor/Activator Protein 1
RPA	Replication Protein A
RQC	RecQ Conserved
spFRET	single pair Förster Resonance Energy Transfer

ssDNA	single-stranded DNA
t-loop	Telomere Loop
TIN2	TRF1 Interacting Nuclear Factor 2
TIRF	Total Internal Reflection Fluorescence
TPP1	TIN2 and POT1 interacting Protein 1
TRF1/2	Telomere Repeat binding Factor 1/2
TRFH	Telomere Repeat binding Factor Homologous Domain
XRCC3	X-Ray Cross Complementing protein 3
WH	Winged Helix
WRN	Werner helicase/exonuclease
WS	Werner Syndrome

TABLES

LIST OF FIGURES

1.0 General Introduction

1.1 TELOMERES AND AGING

1.1.1 Telomere dysfunction and aging

The telomere is composed of repetitive DNA at the ends of linear chromosomes and protects the genome against two endogenous threats to genomic stability(1). First, along with telomerase, the telomere protects the genome against the inability of DNA replication machinery to completely replicate the ends of chromosomes. Secondly, the telomere and the associated shelterin complex of proteins sequester the ends of the chromosomes from the cell's DNA double-strand break repair machinery. The latter function is tightly choreographed by the six shelterin proteins and shelterin interacting accessory proteins such as the Werner Syndrome helicase/exonuclease (WRN) enzyme and p53(1).

Telomere length and stability are related concepts that underlie the link between telomeres on one hand and disease and aging on the other. Mean telomere lengths do decrease with cell division(2), leading to telomeric dysfunction (3). Telomeric dysfunction in turn can lead to senescence and tissue dysfunction, such as in bone marrow(4) and vascular endothelial cells in atherosclerotic blood arteries(5). According to the classical model of telomeres and aging, when the cell has divided a certain number of times—i.e., has reached its Hayflick limit—or has a critically shortened telomere, it must either senescence, apoptose, lengthen the telomeres, or mutate the p53/Rb checkpoints that guard against critically shortened telomeres(reviewed in(6)). The latter pathway is characteristic of the development of a cancerous cell. However, even a cell with adequate telomere lengths may experience genomic instability or premature senescence if the shelterin complex or another important shelterin interacting accessory protein like WRN is not functional(1;7).

1.1.2 Werner Syndrome and telomeric dysfunction as model of aging

Shortened telomeres are correlated with numerous age-related diseases, such as atherosclerosis(5), impaired wound healing(8), type II diabetes(9), reduced hematopoeisis(4;10), and cancer(3). Patients with the progeroid disease Werner Syndrome (WS) manifest many age-related diseases prematurely as a result of lacking functional WRN enzyme(11), and die in their mid-fifties on average(12). While no single model system or biomarker exists for aging, Werner Syndrome provides the opportunity to study numerous tissue-specific age-related pathologies that have made it a common model for aging, with the additional advantage that WS patients have no characteristic anatomic phenotypes (e.g., craniofacial features), such as in Hutchinson-Gilford Progeria (HGP)(11). WS patients thus have a more "wild-type" appearance and their condition is less obvious than other progeria like ataxia-telangiectasia or HGP(13), which suggests that WS may provide a better model for aging in the general population than other progeria.

While WS cells experience normal rates of shortening(14) and the mean telomere lengths are not shorter than in normal cells(15), WS cells experience an accelerated rate of stochastic telomere loss from individual chromatids(16). Moreover, the forced expression of telomerase in human WS fibroblasts suppresses multiple WS phenotypes, including premature senescence(17), sister chromatid telomere loss(18), and chromosomal instability(18). The *Wrn -/-* mouse model displayed no unique phenotype unless the telomerase *Terc* gene was also knocked out and the telomeres were allowed to progressively shorten over several generations(19). Therefore, the value of Werner Syndrome (however imperfect) as a model of aging is linked to telomere dysfunction, but the underlying biochemical and biophysical mechanisms for the telomere-

related genomic instability and their subsequent effects on premature aging have yetto be elucidated.

1.2 THE WRN HELICASE/EXONUCLEASE

1.2.1 WRN structure and function

As a RecQ helicase with an N-terminal exonuclease domain(20) (**Fig. 1.1**), WRN has a complex structure with a rich array of protein-protein and protein-DNA interactions that modulate its activity. As with all RecQ helicases, WRN has two RecA-like domains plus the RecQ-characteristic RecQ Conserved (RQC) and HRDC domains(11). The RQC domain contains a Zn-finger motif that is critical for protein stability in the *Eschericia coli* (*E. coli*) homolog of WRN(21) and also contains the winged helix WH domain, which has both DNA-binding and unwinding roles (22). In the latter case, mutations of the WH domain could eliminate WRN binding to Holliday Junction (HJ) DNA(22). A recent cocrystal of duplex DNA and a WRN fragment revealed that one of the hairpin-motif "wings" separates DNA base pairs(23). The WH domain modulates WRN binding and processivity on DNA substrates(24).

The Helicase and RNAseD C-terminal (HRDC) domain, in contrast to the WH domain, does not bind single-stranded or forked DNA substrates(25). Rather, the HRDC domains seems to be purely a target of protein-protein interactions, such as p53, which interacts with WRN via the HRDC, and inhibits WRN activity on model replication forks and forked double stranded DNA (dsDNA) *in vitro*(26). WRN protein has a rich array of sites that interact with DNA and proteins, including several shelterin and shelterin interacting accessory proteins(26-29). WRN is a RecQ helicase which unwinds DNA in a 3’ to 5’ direction(30), and uniquely has an exonuclease domain that also displays a 3’ to 5’ directionality(31). These domains process DNA

in coordination with each other (**Fig. 1.2**), allowing WRN to displace DNA substrates that the helicase domain could unwind alone(32). As a RecQ helicase, WRN has substrate affinities for DNA structures relevant to genomic maintenance in general, with some being of particular relevance to the telomere, including G-quadruplex (G4) DNA, displacement loops (D-loops) and Holliday Junction (HJ) DNA (30;34).

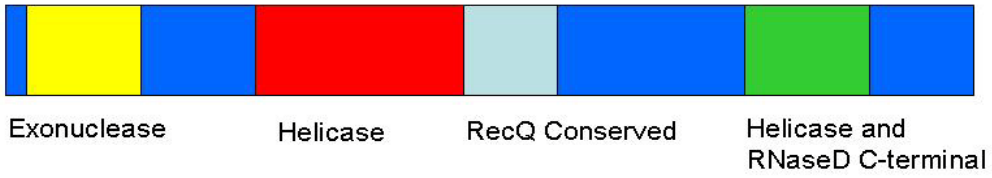

Fig. 1.1 Schematic of the major domains in WRN helicase/exonuclease. The helicase, RecQ Conserved (RQC) and HRDC domains are characteristic of RecQ helicases. The N-terminal exonuclease domain is unique to WRN among the human RecQ helicases(11).

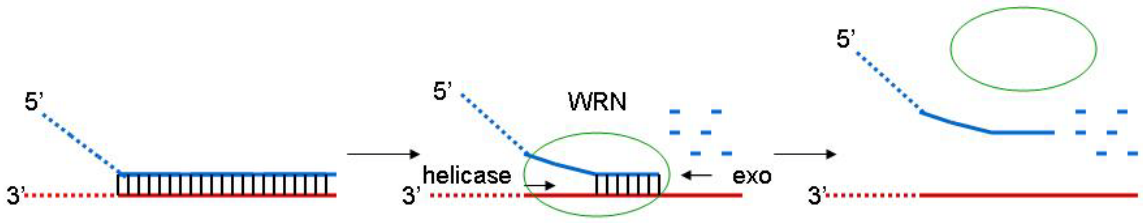

Fig. 1.2 Example of coordinate WRN helicase and exonuclease activity on a DNA fork. The helicase and exonuclease domains attack the substrate in an antiparallel manner in order to displace the exonuclease-shortened blue upper oligonucleotide(see (33)).

1.3 ALTERNATIVE DNA STRUCTURES AT THE TELOMERE

Non-canonical DNA structures (i.e., non B-helical dsDNA) are functionally significant within the telomere, modulating and being modulated by shelterin and shelterin interacting accessory proteins. The human telomere sequence is composed of TTAGGG repeats, first as duplex DNA and as a 3' strand continuing as a single-stranded DNA (ssDNA) extending tail beyond the duplex region (**Fig. 1.3**). The disposition of the ssDNA tail determines the two primary conformations of the telomere.

The first telomere conformation ("closed") is where the DNA end is sequestered by the t-loop/D-loop structure. Bound TRF2 dimerizes and the 3' tail is bent around, creating the lasso-like t-loop and invades the upstream duplex telomeric DNA, displacing the G-rich strand and creating a displacement loop (D-loop) (35). The machinery involved in creating the D-loop structure includes the TRF1/2 shelterin proteins and the homologous recombination proteins RAD51, RAD52, RPA54 and XRCC3(36). The D-loop is a recombination intermediate, and research has shown that telomeres can be cleaved by Holliday Junction resolvases(37), indicating that the D-loop may branch migrate to a full Holliday Junction intermediate as will be discussed in further detail below.

In addition to duplex DNA, the telomere has a 3' tail of single-stranded DNA (ssDNA), composed of 50-100 nucleotides in the case of human cells(38). In the "extended" telomere conformation, the tail is exposed, as it would need to be for replication or extension by telomerase(1).

1.3.1 Mechanisms of WRN function at the telomere

WRN protein has been shown to colocalize to telomeres during S-phase when the genome is replicated(39), and WS cells have prolonged S-phases, suggesting that WRN functions to clear barriers to replication(40).

WRN unwinds G4 DNA (30), and all RecQ helicases share a G4 DNA binding motif in the RecQ Conserved (RQC) domain (41). In addition to prolonged S-phases in WS(40), WRN protein is required for the replication of the G-rich lagging strand of the telomere (18;42), indicating that G4 DNA may be a barrier to replication without WRN protein. Another study found lagging-strand sister telomere loss upon expression of a dominant negative WRN protein (18). WRN has been shown to interact with polymerase δ *in vitro*, allowing it to synthesize DNA through a quadruplex-forming region on the template strand(43).

In addition to G4 DNA, WRN binds to and displaces HJ DNA *in vitro*(30) and WS cells have defects in homologous recombination (HR)-mediated repair of replication forks, which can be suppressed by the simultaneous knockout of RAD51 in WS cells (44). RAD51 binds single stranded DNA (ssDNA) as an initial step in HR; RecQ helicases can suppress RAD51-mediated HR by displacing DNA-bound RAD51(45). Both helicase and exonuclease functions must be present in exogenously expressed WRN in order to process HR intermediates in WS indicating that the two domains work coordinately to resolve HJ DNA(46). While both the helicase(30) and exonuclease domains(31) of WRN are known to process HJ DNA, no experimental data have yet shown coordinated activity between the two on an HJ substrate.

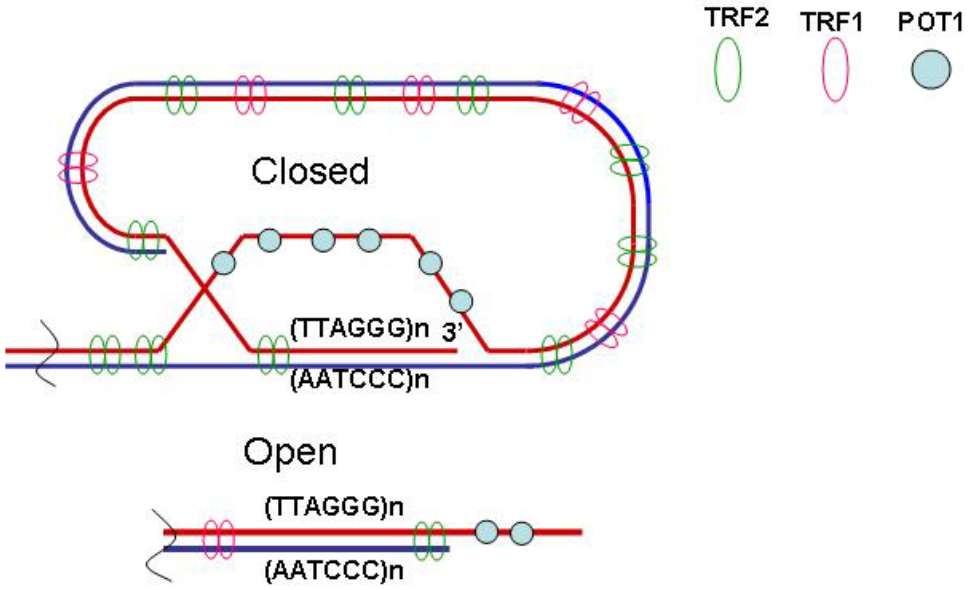

Fig. 1.3 Comparison of the open and closed conformations of the telomere. The three DNA-binding proteins are depicted binding where they can localize in either conformation. Green ellipse, TRF2; red ellipse, TRF1; blue circle, POT1 (1).

1.3.2 Telomeres are protected by the shelterin complex

The telomere is protected by six proteins comprising the shelterin complex(1). Three proteins (POT1, TRF1 and TRF2) directly bind telomeric DNA, and the other three (RAP1, TPP1 and TIN2) coordinate between each other and the DNA-binding shelterin proteins.

The protection of telomeres 1 (POT1) protein binds the G-rich telomeric ssDNA, recognizing the sequence TTAGGGTTAG, thus targeting either the extended 3' tail or the displaced G-rich strand in the D-loop. The N-terminal portion of POT1 has been crystallized, revealing a monomer that binds ssDNA via two oligonucleotide/oligosaccharide binding (OB) folds and a "pocket" that conceals the 3' end of the TTAGGTAG binding sequence(47). This latter feature explains how POT1 has a position-dependent effect on telomerase activity, inhibiting telomerase when the POT1 is at the 3' end (and thus sequestering the telomeric DNA) but promoting telomerase activity when further upstream(48). POT1 also recruits the shelterin protein TPP1 to the telomere. TPP1 adds another layer of control for telomerase activity, as it recruits telomerase to the telomere and increases its processivity(49). In the extended telomere conformation where the ssDNA tail is exposed, the shelterin complex, via POT1 binding the telomeric ssDNA, protects the 3' tail from unwanted ATM and Rad3-related protein (ATR) activation against the ssDNA as a response to stalled replication forks (36).

Telomere Repeat Binding Factors 1 and 2 (TRF1 and TRF2) are both homodimers that bind telomeric duplex DNA with the C-terminal Myb domains(50). Together the two proteins help create the t-loop/D-loop structure. TRF1, because of its highly flexible TRFH dimerization domain(51), can bind duplex telomeric DNA at two sites that are far apart from each other(52), causing the telomeric DNA to bend back on itself(53;54). With TRF1 inducing of t-loop formation, the role of creating the D-loop falls to TRF2. In addition to the Myb domain that is

nearly identical to the TRF1 Myb domain(55), TRF2 possesses an N-terminal basic (B) domain which binds to HJ DNA regardless of sequence(56), catalyzing the formation of the t-loop when the 3' tail is at least six nucleotides long(57). TRF2 recruits RAP1 to the telomere, which confers protection against unwanted non-homologous end joining (NHEJ) at the telomeres(58). NHEJ is a DNA damange repair pathway that repairs double-stranded breaks in DNA, and which causes chromosomal fusions if active against telomeres. TRF1 negatively regulates the telomere length(59). The sixth shelterin protein, TRF2 Interacting Factor 2 (TIN2) (60), binds not only to TRF1 and TRF2(61) but also to POT1, coordinating between ssDNA-binding and duplex-binding shelterin proteins (1).

1.3.3 Alternative DNA structures at the open telomere conformation

The G-rich TTAGGG repeats of the 3' tail can form G-Quadruplex DNA (G4 DNA), a pseudoknot formed by the Hoogsteen base-pairing of four guanine residues into a planar tetrad (**Fig. 1.4A**)(62). In the case of human telomeric DNA, four telomeric repeats will form G4 DNA with three stacked tetrads stabilized by an anion (Fig 1.4B). The TTA repeats act as linkers between the planar tetrads; the precise topology of the linkers can vary greatly depending on anions(63;64), crowding conditions, and the presence of flanking DNA(65), among other factors. G4 DNA can form at the 3' tail of the telomere in the telomere open conformation or in the displaced G-rich DNA in the D-loop in the telomere closed conformation. Telomeric G4 DNA been detected in *S. lemnae* telomeres using immunofluorescence(66). Outside of the telomeres, G4 DNA has been observed with electron microscopy in the non-coding G-rich strands of transcription bubbles(67)

Structural studies of G4 DNA have focused on individual G4s, including all Protein Data Bank (PDB) files from Nuclear Magnetic Resonance (NMR) and X-ray crystallography (e.g., (62;63;68-71)). However, since the human telomere may have up to 30 telomeric repeats(1), multiple quadruplexes may fold, and might even fold "stacked" structures of multiple tetrads(72) that form a rigid structure similar to the guanine-wire (G-wire), which is a long nanowire composed of intermolecular G-tetrads(73).

The structure of G quadruplexes on long telomeric tails has important implications for the structure of the chromosome. A recent study demonstrated that G4 DNA formation promotes telomerase activity in yeast (74).Telomere End Binding Protein β (TEBPβ, the homolog of the TPP1) facilitates the formation of G4 DNA in *Oxytricha nova*(75). While the function does not exist in human TPP1, quadruplexes are functionally vital in gene promoter sites(76) and as a telomerase promoter in fission yeast (74), suggesting that quadruplexes may have both desirable as well as harmful effects on the telomeric homeostasis.

While no quantitative structural studies of long telomeric tails exist, bulk studies have suggested that individual quadruplexes form as "beads on a string" (**Fig. 1.5**) on long telomeric tails rather than the rigid stacked formation. Melting studies demonstrated that telomeric tails with telomeric repeats in integer multiples of four (e.g., 8, 12, 16) had comparable thermodynamic stabilities, with non-integer multiples being less stable(72). However, another study that used potassium phosphate buffer instead of Tris found that longer telomeres were progressively less stable due to enthalpic penalties (77).

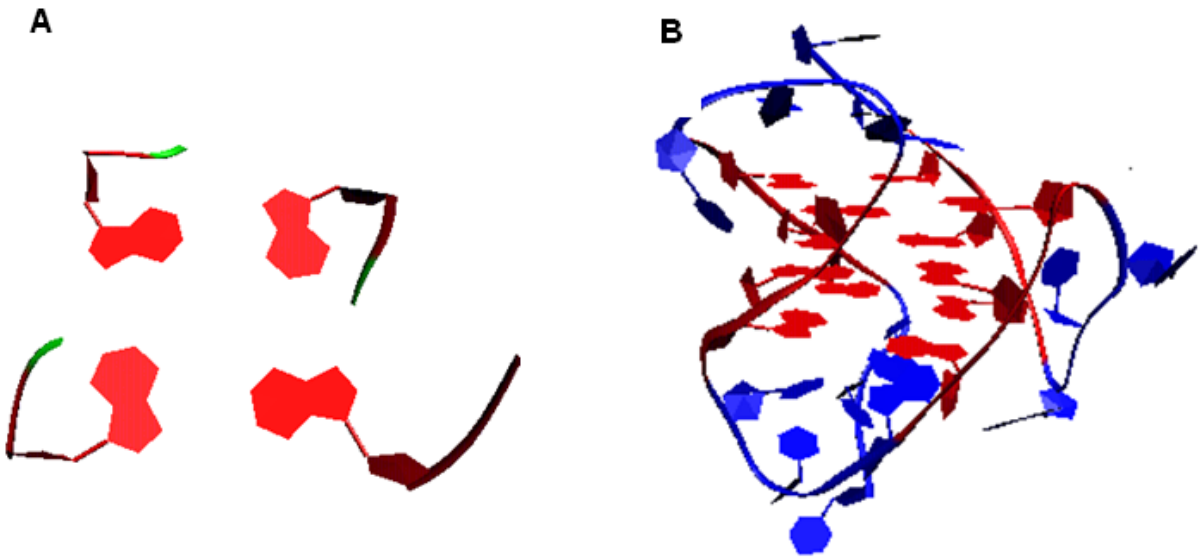

Fig. 1.4 Structure of human telomeric G4 DNA. **A,** an individual tetrad of four guanines bound by Hoogsteen base-pairing with each other. **B,** the human G quadruplex structure elucidated by NMR, with the flat, planar guanine tetrads visible in the middle in red. Adenine and thymine loop residues are in blue. Both subfigures are from structure 2JPZ (65), accessed from the Protein Data Bank (www.pdb.org) and depicted using VMD (University of Illinois Urbana, Illinois).

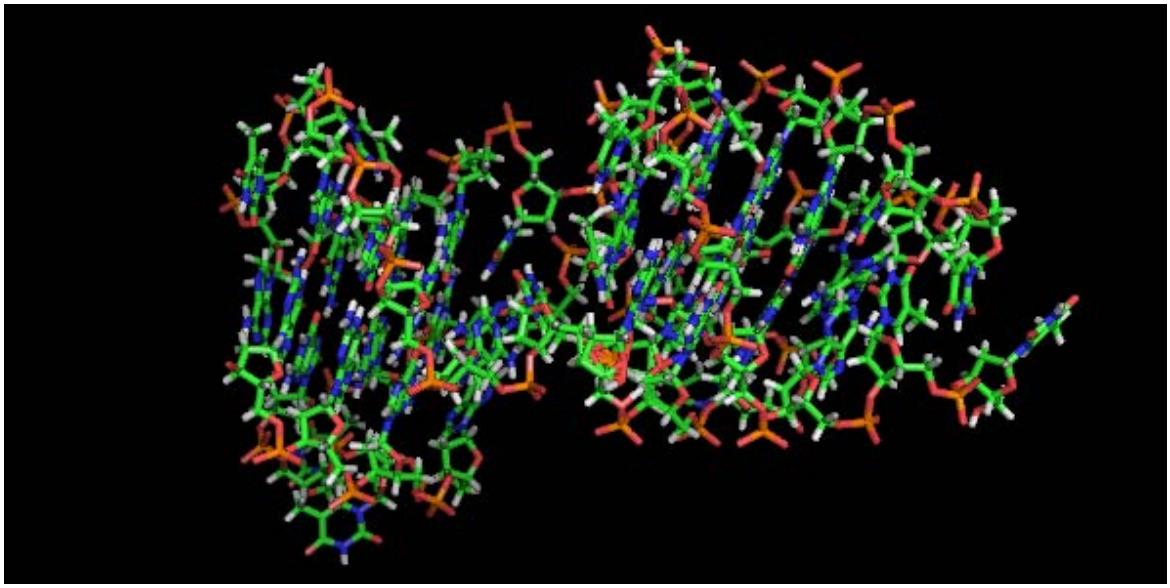

Fig. 1.5. Proposed model of a beads on a string arrangement of two quadruplexes. Two copies of the 2JPZ structures from **Fig. 1.4** were linked together using Pymol after removing the two 3' thymines from the upstream quadruplex. Note the two peaks separated by the central gap between the quadruplexes.

The UV melting data demonstrating the stability bonus of having a multiple of four telomeric repeats also suggested that quadruplexes fold in such a way that they maximize the number of quadruplexes formed (e.g., three quadruplexes on 12 telomeric repeats) (78). However, recent circular dichroism (CD) studies on a $(TTAGGG)_{13}$ substrate showed that titrations of a quadruplex-loop binding molecule indicated that only two quadruplexes formed(79). Since these values are only averages of bulk populations, the makeup of a molecular population with G4 forming potential is still unknown, although they are functionally relevant to the extended or open telomeric tail telomere conformation.

1.3.4 Alternative DNA structures at the closed telomere conformation

A linear chromosome must hide its ends from the DNA repair machinery of the cell. The 3' ssDNA tail provides an elegant solution to this problem by invading the upstream duplex telomeric DNA and creating the D-loop portion of the t-loop/D-loop structure. This strand invasion is catalyzed by the shelterin protein TRF2, a homodimer that has two binding sites on each monomer. The N-terminal B domain binds HJ DNA centers in a sequence independent manner(56), and the C-terminal Myb domain specifically binds duplex telomeric DNA(55). TRF2, but not TRF1, catalyzes strand invasion via the N-terminal domain by inducing positive supercoils in the duplex telomeric DNA(57). Similar to HJs, a replication fork may stall, causing the newly synthesized strands to anneal to each other, peeling away from the templace strands. This regressed replication fork or “chickenfoot” structure is a possible site for TRF2 binding, as telomeric replication forks are prone to regression (80).

TRF2 maintains telomeric integrity via three chief pathways, the first being the suppression of the ataxia telangiectasia mutated protein (ATM) pathway of double strand break

repair by binding the RAP1 shelterin protein. TRF2 itself does not suppress the ATM pathway but rather mediates ATM suppression by recruiting RAP1 to the telomere(81). TRF2 also protects the telomere D-loop/t-loop structure from HJ cleavage via its N-terminal, Basic (B) Domain, which bind HJ centers regardless of sequence (82). If a deletion mutation of TRF2 missing the B domain (TRF2ΔB) is expressed in a cell, telomeres undergo sudden shortening with telomeric circles appearing in both 2D electrophoresis and electron micrographs(37). The phenotype is suppressed when the X-Ray Cross Complentation 3 (XRCC3) homologous recombination protein is deleted, indicating that HJ cleavage is the cause for sudden telomere shortening(37). In WS cells, however, TRF2ΔB expression does not cause sudden telomere shortening from HJ cleavage activity, possibly because TRF2 prevents WRN from catalyzing the branch migration of a D-loop into a full HJ that can be cleaved(83).

The mechanism of TRF2 in stabilizing HJ DNA at the telomere has been partly characterized by biochemists. Poulet and colleagues have shown that the B domain of TRF2 is necessary to protect an artificial HJ against the HJ cleaving enzyme XPG-like endonuclease (GEN1) *in vitro*(84). No similar biochemical studies, however, have examined how TRF2 modulates WRN activity on HJ DNA in vitro. Since TRF2 stimulates WRN activity on telomeric duplex DNA, it is not obvious that TRF2 necessarily would inhibit WRN-mediated displacement of HJ DNA.

WRN is not the only helicase that can be deleterious if it is not tightly regulated by the shelterin complex. In fission yeast, the WRN homolog Rqh1 causes telomere loss if the TRF1/2 homolog Taz1 and the DNA-binding protein RPA is deficient(85). Mutating SUMOylation sites in Rqh1 in the absence of functional Taz1 leads to telomere loss and hyperrecombination and inter-chromosomal telomeric entanglements(86). Elucidating the mechanism of WRN regulation

by TRF2 would provide an important contribution to the body of knowledge on how telomeric DNA structures are maintained by the shelterin complex.

1.4 SINGLE MOLECULE ATOMIC FORCE MICROSCOPY IMAGING

Atomic Force Microscopy (AFM) uses minute deflections in a thin needle (tip radius is under 10 nm) to record a topographical map of molecules deposited on an ultraflat surface. The two primary modes of AFM are contact or tapping (reviewed in (87;88)). In the former, the tip is kept at a certain height above the surface. When the tip is deflected by a surface feature, the feedback mechanism in the AFM moves the tip so that it will be at the preset height. These deflections are used to calculate the height of the surface at a given x-y coordinate to build up an image of the surface. In the tapping mode, the tip oscillates at a set amplitude of vibration, which is altered when the tip approaches an object. A feedback loop similarly adjusts the height of the tip to maintain the amplitude of vibration. The modulations in the tip's frequency are then used to reconstruct the proximity, and thus, the height of an object. The tapping mode is less direct and thus, preferred for biomolecules, which can be distorted by direct contact. AFM may be done on a dry sample in the air, or it may be done in a fluid-filled container, the latter being more complicated but permitting one to study live cells and protein dynamics(87). This introduction will focus on the air AFM method.

Because AFM can identify individual molecules at the nanometer scale, it is excellent at distinguishing diverse molecular conformations. AFM has been successfully employed to quantify single stranded, duplex and triple-stranded DNA(89), Holliday Junctions(90;91) and processing of DNA by DNA repair enzymes(92;93). While there have been AFM studies of G4 DNA, these focused on non-telomeric sequences(94) and single telomeric quadruplexes(95), and

the limited studies that did involve telomeric DNA were not quantitative(96;97). AFM is well suited for studies on long telomeric ssDNA substrates, where multiple G4 conformations and POT1 binding registers are possible.

While the bulk data on G4 DNA are valuable, quadruplexes are a difficult subject for bulk studies since they are dynamic structures that may spontaneously fold and unfold(64). Moreover, the topology of the loops in between G4 tetrads and their orientations are highly dependent on the concentration of the DNA, the cation used and its concentration (63;64), flanking sequences(65), and the presence or absence of crowding agents(95;98), to name some primary examples.

Important data on single quadruplexes have come from single molecule Foerster Resonance Energy Transfer (FRET), demonstrating the role of cations and temperature on quadruplex stability and loop orientation(64). FRET employs a donor and an acceptor dye pair where the excitation spectrum of the acceptor overlaps with the emission spectrum of the donor. Therefore excitation of the donor dye will result in the excitation and emission of the acceptor by an induced dipole, with the amplitude of the acceptor dye being in inverse proportion of the distance from the donor, creating a molecular ruler in real time(99). However, a longer substrate with multiple possible folding G4 conformations does not lend itself to FRET studies since one would require a prohibitive number of substrates with different dye pairs. Atomic force microscopy has the advantage of studying individual molecules like FRET but by visualizing the shapes of individual molecules (rather than the distance between two dyes molecules), AFM produces a more detailed picture of DNA folding than FRET. Unlike bulks studies, AFM does not measure an average value but permits a whole range of conformations to be individually inspected and quantified.

1.5 STATEMENT OF THE PROBLEM

Telomeric instability is correlated with age-related diseases and tissue dysfunction (3). The stability of a telomere is maintained by a shelterin complex of six proteins at the telomere plus numerous shelterin and telomere accessory proteins like WRN and p53 (1). Many of the accessory proteins can be deleterious to telomeres if they are not regulated by the shelterin proteins. In addition to the proteins interactions, non-canonical DNA structures like HJs and G4 DNA add further layers of complexity in telomere function and dysfunction. We hypothesize that alternate structures of DNA exist in dynamic equilibrium in telomeric DNA and the equilibrium is modulated by WRN helicase/exonuclease in cooperation with telomere-binding proteins.

The first goal is to determine the ability of WRN and POT1 to modulate G4 formation on long physiologically relevant telomeric ssDNA tails. Most studies have focused on individual quadruplexes, but physiologically relevant telomeric DNA lengths can fold into multiple quadruplex arrangements. Furthermore, studies on long telomeric substrates have largely consisted of bulk experiments, measuring the mean values of what could be a highly heterogeneous population. The second goal is to elucidate the modulation of Holliday-Junction dynamics by TRF2 and WRN. We developed fluorophore imagery to track WRN processing of HJ DNA containing differing arrangements of telomeric repeats with and without TRF2 and/or the TRF2-binding protein RAP1.

1.6 OUTLINE OF THESIS

Chapter 1 is a general introduction to telomere homeostasis and the value of the progeroid disease Werner Syndome as a model telomeric pathways of aging. The structure and function of the Werner helicase/exonuclease (WRN) will be introduced, and then discussed in the context of modulating DNA structures in the open and closed conformations of the telomere. The shelterin complex of telomeric proteins will be then introduced and their interactions with each other and their modulation of DNA structures will be discussed. Finally, the Atomic Force Microscope and its role in single molecule biology will be reviewed.

Chapter 2 describes the characterization of G quadruplex (G4 DNA) formation on physiologically realistic telomeric tails by AFM imagery, and the destabilization of G4 DNA by the POT1 proteins.

Chapter 3 discusses the elucidation of the role of TRF2 in protecting telomeric Holliday Junctions against unwanted WRN helicase/exonuclease activity. The mechanism of HJ displacement by WRN is analyzed and discussed in the context of interactions with TRF2 at the telomere.

Chapter 4 provides an overview of the experimental results and an analysis of some of the limitations of the AFM system in visualizing the G4 DNA system in Chapter 2. Further avenues of research using both AFM and single molecule Förster Resonance Energy Transfer (FRET) are discussed.

2.0 STRUCTURE OF G-QUADRUPLEX ASSEMBLIES ON LONG TELOMERIC SINGLE STRANDED MOLECULES IS MODULATED BY POT1

This chapter is currently being prepared as a manuscript by the following authors: Hong Wang[†], Gerald J. Nora[†], Harshad Ghodke, Bennett Van Houten, and Patricia L. Opresko.

[†]These authors contributed equally to the paper. H.W. and G.J.N. proposed the project, designed and performed most of the experiments, analyzed the data, and wrote the manuscript; H.G. performed some of the AFM experiments; P.L.O. supervised the project, interpreted the data and wrote the manuscript.

2.1 INTRODUCTION

Cells with linear chromosomes must solve two problems: the progressive lagging-strand shortening with each cycle of DNA replication and the need to protect the ends of linear chromosomes from unwanted DNA damage responses(1). The telomere is a solution to both these problems and stands at the junction between aging, genomic stability and cancer. Telomeres are composed of the "shelterin complex" of proteins and TTAGGG repeats of duplex DNA along with an ssDNA overhang or "tail" of 50-500 nucleotides(1). The ssDNA tail can fold into G-quadruplex (G4) DNA, which consists of three tetrads of four guanines that form Hoogsteen base pairs with each other (**Fig. 2.1A**). These tetrads are in a square-planar conformation and are stacked atop one another with the TTA sequences forming linker loops(100;101). The formation of G4 DNA has been shown to inhibit the telomere-lengthening enzyme complex telomerase *in vitro*(102), although a recent *in vivo* study of *S. cerevisiae* telomerase found that G4 DNA can promote the activity of yeast telomerase(74).

The folding of G4 DNA is inhibited when the DNA is bound by the protection of telomeres 1 (POT1) protein(103), which is part of the shelterin complex of proteins that bind, protect and regulate telomere structure and function. The mechanisms underlying the POT1 modulation of intra-molecular G4 DNA assembly on physiologically relevant telomeric ssDNA tails in which multiple G4 DNA structures are possible, are unknown. Studying POT1 loading on physiological telomeric tails is complicated by a lack of data on G4 DNA formation on long ssDNA strands. X-ray crystallographic and NMR studies of G4 DNA have focused on individual quadruplexes formed from four TTAGGG repeats(62;63;68-71). Since the telomeric tail may have upwards of 30 such repeats(1), it is important to understand the arrangements of quadruplexes, and how many quadruplexes can form on longer telomeric ssDNA. Results from

thermal melting assays support the hypothesis that longer telomeric repeats form a beads-on-a-string G4 assembly in which individual quadruplexes are separated from each other by a TTA linker (**Fig. 2.1B**)(78), although some data and extrapolations from an NMR structure support a "stacked" arrangement of quadruplexes(104;105).

Atomic force microscopy (AFM) offers a single-molecule approach that allows one to examine distinct nucleic acid structures (single-, double- and triple-stranded) and their distribution within a heterogeneous population(106;107). In addition, previous AFM studies established the visualization of single human telomeric G4 DNA by AFM(95). However, the quantitative distribution of various quadruplex numbers and arrangement ensembles of individual molecules within a potentially very heterogenous population of long single-stranded telomeric molecules have not been addressed. Even more importantly, POT1 coats the 3' ssDNA tail of the telomere(108;109), and POT1 can greatly influence the distribution of G4 structures within a population of telomeric DNA. However, modulation of G4 folding by POT1 on physiologically relevant telomeric tails has not been investigated at the single-molecule level. AFM has been used extensively in studying protein-DNA interactions(110;111), validating its application in visualization of the modulation of the telomeric tail structure by POT1.

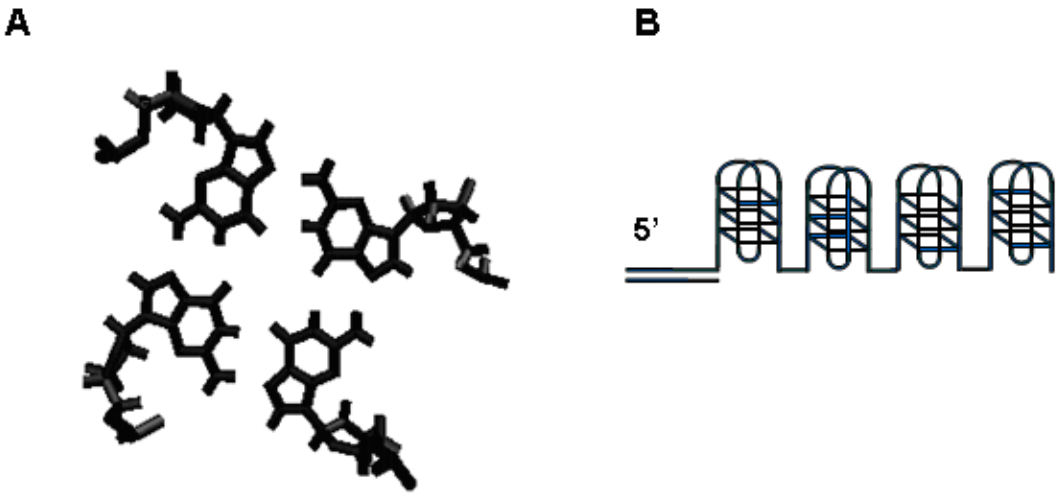

Fig. 2.1 Base pairing in G4 DNA and schematic of beads-on-a-string model. (**A**) Planar tetrad of guanines bound by Hoogsteen base pairing. Image created on VMD (Urbana, Illinois) using the Protein Bank 2JPZ structure(104). (**B**) Schematic illustration of the beads-on-a-string model(62). In this model, long single stranded telomeric DNA form a beads-on-a-string G4 assembly in which individual quadruplexes are connected with each other by an ssDNA linker.

First, to study the formation of G4 DNA on realistic telomeric tails, we performed single-molecule AFM imaging of defined DNA substrates with a duplex stem followed by single stranded TTAGGG repeats (4, 8 or 16), and conducted detailed quantitative analysis of the length, height, and volume of the G4 structures. At physiological salt concentrations, 55% of $(TTAGGG)_{16}$ molecules formed only two G4 structures consistent with a beads-on-a-string conformation, instead of the maximum of four that would occur if all the TTAGGG repeats folded into quadruplexes. Secondly, we found that POT1 is a monomer with a height differing from G4 DNA under AFM imaging conditions. POT1 can coexist with G4 DNA on the same 3' tail and shift the population distribution toward molecules that have half the number of G4 units as substrates without POT1, or that are completely unfolded. These novel results provide a structural basis for the dynamic interplay between G4 DNA formation/disruption and POT1 binding at the telomeric 3'tails.

2.2 RESULTS

2.2.1 Long telomeric tails are underfolded

Prior to studying POT1 modulation of G4 DNA on physiological telomere tails, we set out to elucidate G4 DNA structures on these molecules in the absence of POT1. Previous AFM studies of G4 DNA used either short telomeric sequence (four repeats), 3' tails of unknown lengths, or did not provide quantitative or distribution analysis of the images(112-114). Consequently, detailed information regarding the distribution and types of conformations of physiological telomeric tails required further study. We designed a series of defined DNA substrates which have a 34 bp duplex stem at the 5' end followed by a 3' ssDNA overhang of 4, 8 or 16 TTAGGG repeats (Tel4, Tel8 and Tel16, respectively, **Table 2.1**). Tel4, Tel8 and Tel16 can potentially

form a maximum of 1, 2, and 4 G4 units, respectively. We reasoned that comparison of G4 structures formed on these substrates as visualized through AFM imaging, would provide quantitative information regarding the number of G4 units present on each molecule. AFM field view images and surface plots of Tel4, Tel8 and Tel16 show that all three telomeric substrates formed structures with heights between 1 and 2 nm (**Figs. 2.2,2.3, and 2.8**), which were not observed in images of duplex DNA or an ssDNA substrate which lacks G4 forming sequences (**Figs. 2.8 and 2.2E**, respectively). The heights of peaks observed for the Tel4, Tel8 and Tel16 substrates are consistent with the height measurements from previous AFM studies of G4 DNA(115).

Since the height of G4 structures (average height, 1.32 ± 0.22 nm) exceeds the height of duplex DNA (0.61 ± 0.1 nm, **Fig. 1D**), we measured the length of G4 structures at 1 nm height (**Fig. 3**) to compare the amount of G4 DNA formed on Tel4, Tel8 and Tel16. In addition, we measured the volumes of G4 DNA derived from the AFM images. AFM-derived volume has been used extensively to study the oligomerization states of globular proteins and protein-protein interactions(110;116). The lengths (at 1 nm height) and volumes of Tel4, Tel8 and Tel16 are shown in **Figs. 2.2D** and **2.8**, respectively. The mean lengths of G4 DNA at 1 nm height of Tel4 and Tel8 are 12.5 (± 1 nm, 3 depositions) and 15 (± 3 nm, 3 depositions) nm, respectively. The mean length of Tel8 is slightly longer than Tel4, but is not doubled as would have been expected if Tel8 formed the maximum possible number of quadruplexes. All Tel16 molecules in the AFM images displayed peak heights greater than 1 nm (**Fig. 2.3C**). The mean length of Tel16 (22 ± 2 nm, 3 depositions) is only approximately twice that of Tel4, even though Tel16 could theoretically form a maximum of four repeats

Figure 2.2

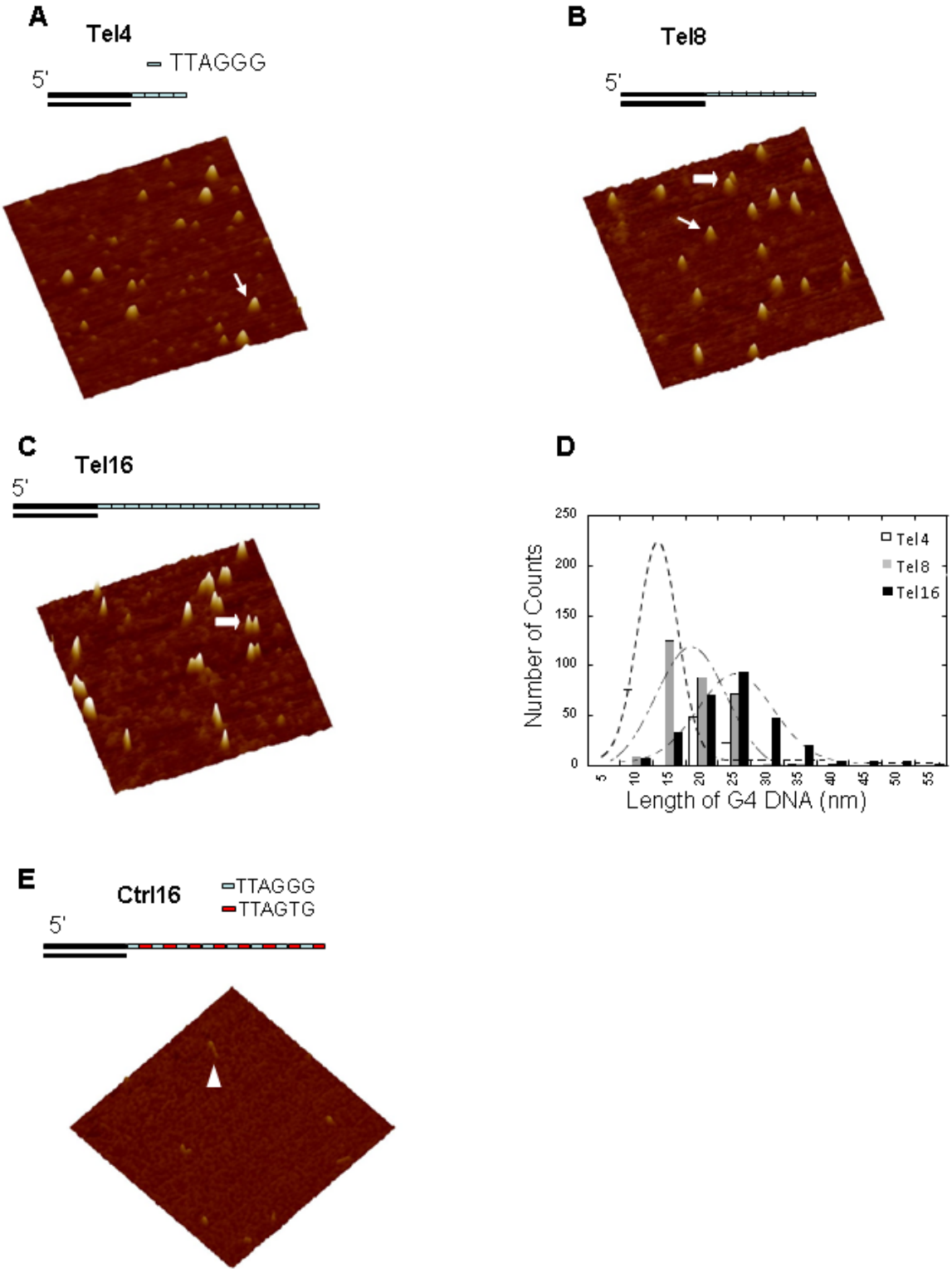
A
Tel4
TTAGGG
5'
B
Tel8
5'
C
Tel16
5'
D
Number of Counts
Length of G4 DNA (nm)
Tel4
Tel8
Tel16
E
Ctrl16
TTAGGG
TTAGTG
5'

Fig. 2.2 Quantification of the number of G4 structures formed on Tel4, Tel8 and Tel16. (**A-C)** Schematic drawings and representative AFM surface plots of Tel4 (**A**), Tel8 (**B**) and Tel16 (**C**) DNA substrates. Thin arrows point to single G4 structures and wide arrows denote two distinct G4 structures on individual Tel8 or Tel16 molecules. The sequences of the substrates are shown in Supplementary Table 1. All DNA substrates were incubated in a buffer containing 150 mM KCl and deposited at 500 nM concentration (see **Methods**). Minor particles in **A** are likely contaminants in the Tel4 preparation, such as acrylamide from the gel purification by the manufacture, rather than unfolded molecules since these images differ from unfolded Ctrl16 structures. (**D**) The histogram of G4 length (cross-section at 1 nm height) measured from AFM images of Tel4 (open bars, n=374 molecules), Tel8 (gray bars, n=299 molecules) and Tel16 (black bars, n=288 molecules). The dotted lines are Gaussian fits to the data. (**E)** A representative AFM surface plot of Ctrl16 DNA, which contains 8 TTAGGGTTAGTG repeats (**Table 2.1**) and does not form G4 structures. The triangle points to an individual Ctrl16 molecule. All images are 500 nm × 500 nm at 2 nm height scale.

Figure 2.3

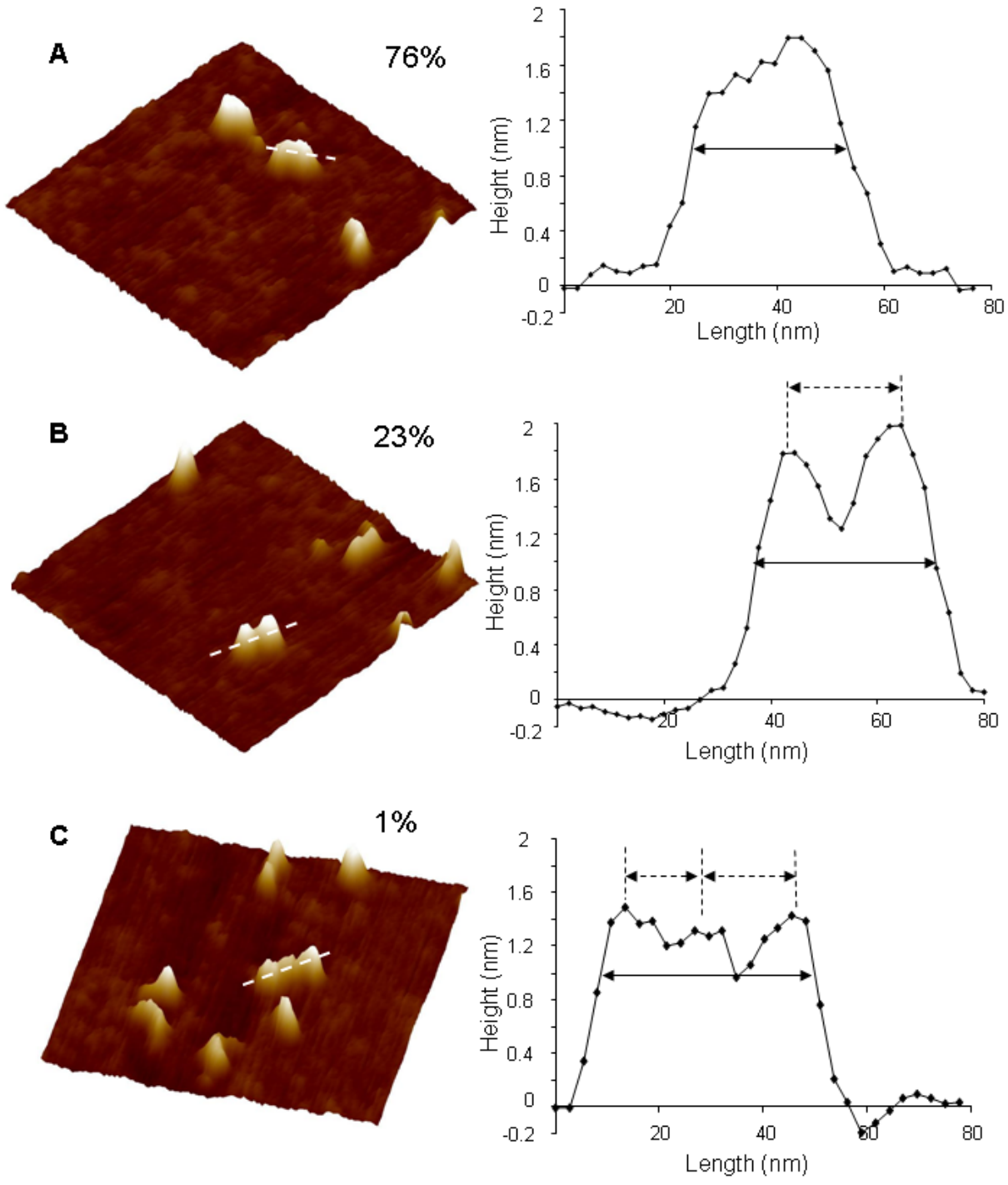
A
76%
Height (nm)
Length (nm)
B
23%
Height (nm)
Length (nm)
C
1%
Height (nm)
Length (nm)

Fig. 2.3 Subpopulation of Tel16 molecules resembles beads-on-a-string. A representative AFM image (left panel) and sectional analysis (right panel) of a Tel16 molecule in which individual G4 structures cannot be resolved (**A**) and a Tel16 molecule which contains two distinct peaks (**B**). (**C**) A representative image of a Tel16 molecule with three distinct peaks. The AFM images are 250 nm × 250 nm at 2 nm height scale. The white lines in AFM images indicate the lines drawn for section analysis. The solid lines with arrows in the section analysis indicate the length of G4 measured at 1 nm height; the dashed lines with arrows indicate the interpeak distances. The number at the top right corner of each image indicates the percentage of each type of molecules observed in the total population of Tel16 molecules.

To investigate the mechanism underlying the underfolding (i.e. formation of less than the maximum potential number of quadruplexes) of Tel8 and Tel16, we constructed a first-principles combinatoric model (see **Methods**) considering each telomeric repeat as a lattice point which can either be extended or folded into G4 DNA (**Fig. 2.4A**). The model shows that the formation of a single G4 in Tel8 is nearly five times more probable than two G4 structures (calculated in **Methods**). For Tel16, the most striking insight from the combinatoric model is that formation of four G4 structures on Tel16 is a rare event, which is consistent with our experimental data. In addition, the folding of two quadruplexes was the most probable conformation (**Fig. 4B**), but three quadruplexes were almost as probable as two. This did not fit the normalized experimental data in which the lengths of G4 regions from Fig. 1 were divided by the mean G4 length (12.5 nm) from the Tel4 data (**Fig. 2.4B**). A previous study reported indirect evidence that an oligonucleotide with 13 telomeric repeats formed only two quadruplexes(79). Assessing whether or not the combinatoric model is consistent with the results using 13 telomeric repeats, we calculated the probability distributions for DNA containing 13, 14 and 15 repeats.

We found that both substrates with 13 and 14 repeats both exhibited maxima for two G4s, but for a molecule with 15 repeats, three G4s was highly probable as well (**Fig. 2.4B**).

2.2.2 Long telomeric tails form beads-on-a-string structures

Different models have been proposed to describe the intra-molecular assembly of multiple G4 units on long telomeric ssDNA(78;117;118). In a beads-on-a-string model, two G4 units are connected by one linker without stacking interactions between the units (**Fig. 2.1B**). In the stacking model, every G4 unit stacks onto adjacent G4, with residues on the TTA loops interacting with each other(78;105;119).

Figure 4

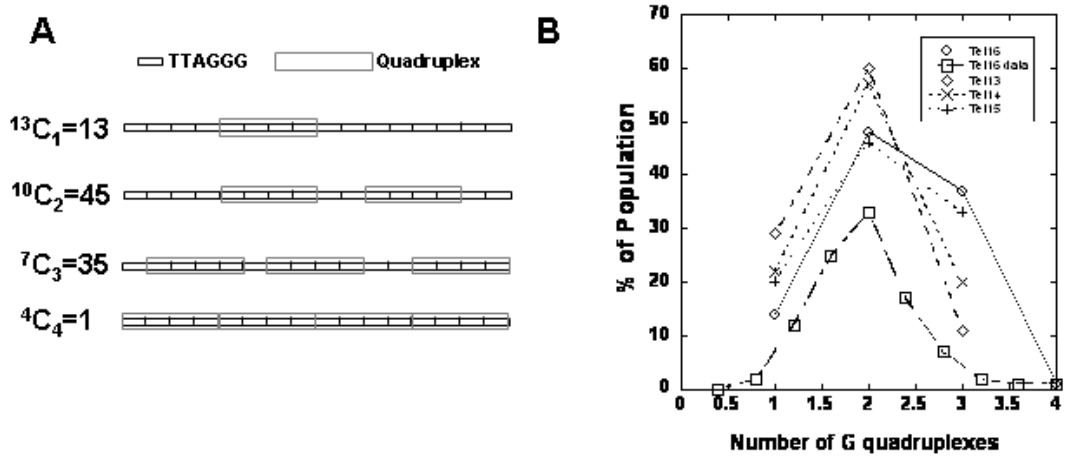

Fig. 2.4 Longer telomeric tails rarely form the maximum potential number of quadruplexes. (**A)** Schematic examples and calculations of the number of possible arrangements of G4 DNA on Tel16 (see **Methods** for equation). (**B**) Probability of forming 1, 2, 3 and 4 G4 structures on substrates with 13, 14, 15, and 16 (Tel13-16, respectively) based on the combinatoric calculations detailed in **Methods**. The Tel16 data curve is based on the length of G4 regions on Tel16 molecules normalized using the length of single G4 measured from AFM images of Tel4 (12.5 nm).

Among all the Tel16 molecules observed, 23% displayed two distinct peaks in the AFM images (**Figs. 3B**). The mean interpeak distance of Tel16 molecules with two distinct peaks is 20 nm, which corresponds to ~7 TTAGGG repeats between the individual quadruplexes (**Fig. 2.12**). In the AFM images of Tel16 molecules, a small population (1%) of molecules exhibited as many as three distinct peaks (**Fig. 2.3C**). The assembly of multiple defined peaks resembles individual beads-on-a-string. It is worth noting that due to limitations in the AFM resolution, results from AFM imaging could underestimate the number of Tel16 molecules forming beads-on-a-string structure (see calculations in **Chapter 4.1**).

To further differentiate between the beads-on-a-string and the stacking models, we imaged G-wires which are long complexes of highly ordered self-assembly of inter-molecular G4 units (**Fig. 2.5A**). G-wires are notable for reliably creating long, uniform quadruplectic structures with heights greater than 1 nm(120). AFM images of G-wires formed by the short oligonucleotides $G_4T_2G_4$ are shown in Figs. 5B and 5C. Since the G-wires involve stacking of the adjacent G4 units, as expected, regular well separated peaks were not apparent in the AFM images as expected, even for G-wires that were the same length as Tel16 molecules (**Fig. 2.5C**,**D**). In addition, G-wires exhibited a statistically significant ($p < 0.008$) greater average height (1.63 ± 0.17 nm) compared to the Tel16 structures (1.32 ± 0.22 nm) (**Fig. 2.6B**). These results suggest that the G-wires appeared to be more rigid possibly due to the direct stacking interactions between adjacent G4 units, which lead to less compression by the mechanical AFM imaging process. The distinctly different structure and stiffness of the G-wires compared to the Tel16 molecules revealed by AFM imaging suggest that G4 structures on Tel16 molecules are inconsistent with a stacked model of multiple G4 units.

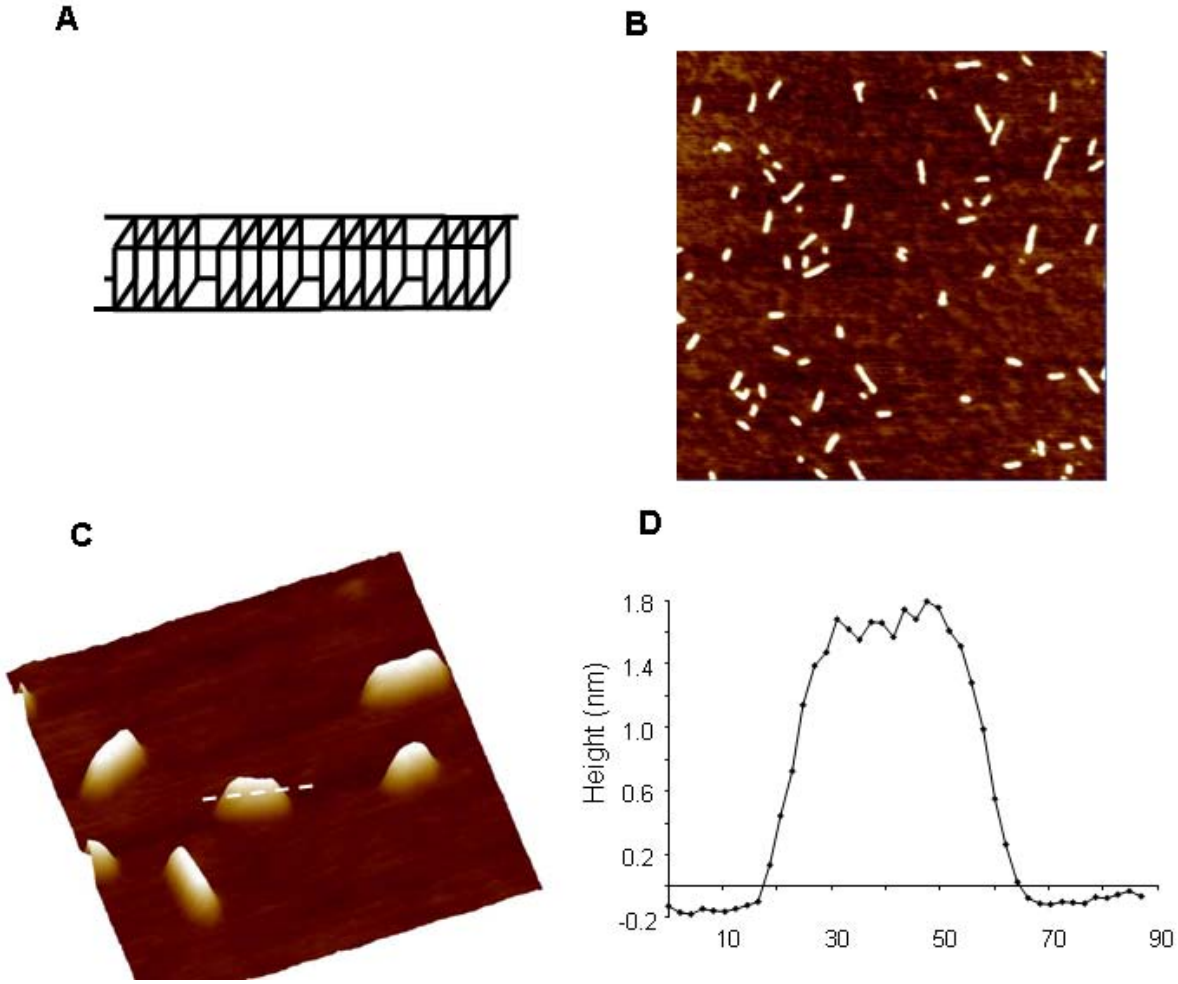

Fig. 2.5 AFM imaging of G-wires reveals a smooth surface without distinct peaks. (**A**) Proposed model for G-wire formation. (**B**) An AFM field view image of G-wires. The image is 1 μm× 1 μm at 2 nm height scale. (**C**) An AFM surface plot of G-wires. The AFM surface plot is 250 nm × 250 nm at 2 nm height scale. The white line denotes the line for section analysis. (**D**) Section analysis of G-wire highlighted in (**C**).

2.2.3 Oligomeric state of POT1

A key issue in understanding the mechanism of action by POT1 is its oligomeric state. Despite evidence showing a monomeric state for the N-terminal domains of human POT1(121), information on the oligomeric state of full length human POT1 proteins was lacking. To evaluate the oligomeric state of full length POT1 we measured the volume of POT1 in AFM images compared to other known proteins of various sizes. AFM-derived volumes of proteins can be correlated to their molecular masses, permitting determination of oligomeric states (see **Methods**) and protein-protein interactions(122). Purified POT1 protein after removal of the GST-tag appeared as monodispersed particles in the AFM images (**Fig. 2.7A**).The distributionof the calculated AFM-derived volumes of POT1 is Gaussian and centered at ~22 nm^3, which is consistent with the expected value for a POT1 monomer based on our calibration curve for globular proteins (**Fig. 2.14**). These results demonstrate that POT1 exists as a monomer in solution under the AFM imaging conditions tested. In contrast, AFM images of GST tagged POT1 protein (GST-POT1) revealed particles larger than untagged POT1 protein (**Fig. 2.8A**). Statistical analysis of the AFM-derived volume of GST-POT1 indicated a distribution consistent with the presence of GST-POT1 dimers and tetramers (**Fig. 2.8B**). These results demonstrate that the GST tag can artificially force POT1 to form dimers and higher order oligomers. Therefore, only untagged POT1 was used in all the imaging experiments with the DNA substrates.

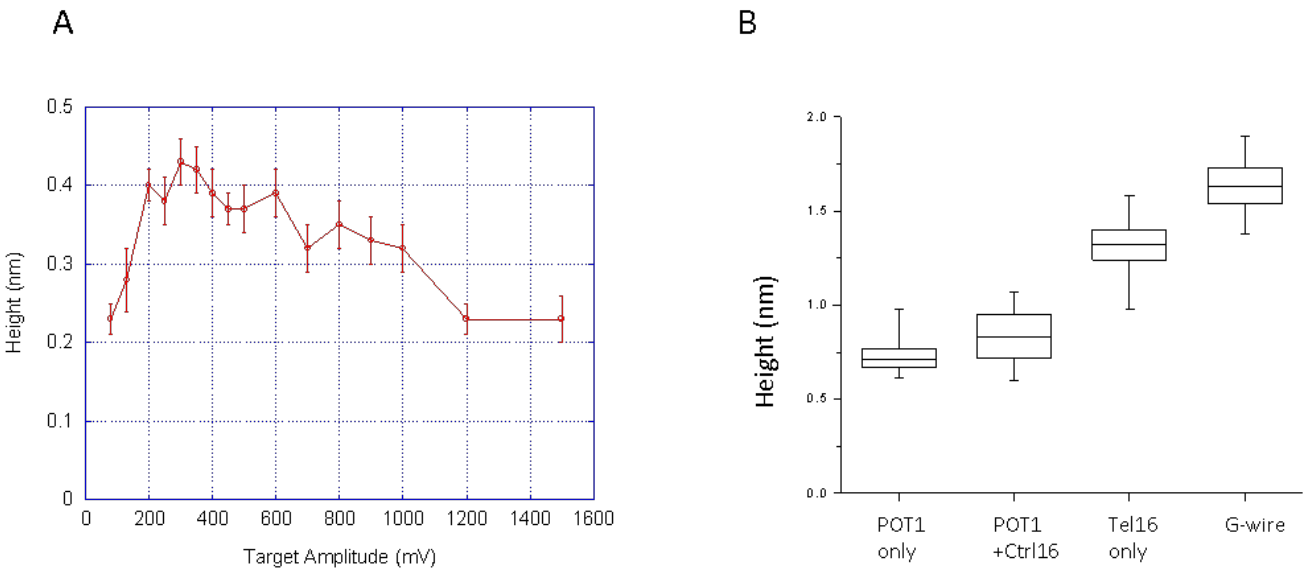

Fig. 2.6 Height as a function of amplitude and substrate. (**A**) The height of hemagglutinin (HA) antibody in AFM images as a function of target amplitude. Monoclonal HA antibody (50 nM, Covance) was deposited onto mica and imaged under standard conditions (**Methods**). To generate each data point, the same area of the mica surface was imaged and the average height of 60 HA antibody molecules was calculated, for each value of the target amplitude for the Nanoscope V. Each error bar represents one standard deviation in the height. It is worth noting that the measured height is maximum at 300 mV. This is because in the low target amplitude regime, the background noise tends to be high leading to a lower corrected height whereas at higher target amplitudes, the force exerted on the molecules is sufficiently large to cause significant compression of the molecule, leading to a reduced measurement of the height. (**B**) Box plot of AFM average heights of all the molecules in AFM images of POT1 only (n=1262 molecules), POT1 in the presence of Ctrl16 (n=304 molecules), Tel16 only (n=235 molecules) and G-wire (n=444 molecules). The image background height is at 0.3 ± 0.05 nm. The height measurements were done using Image SXM software as described in **Methods**. Each box

represents the interquartile range (25^{th} to 75^{th} percentiles). The horizontal bars in the boxes represent the medians. The vertical bars represent the range of values included in the statistics, from the maximal to the minimal.

2.2.4 POT1 binding competes with G4 formation

To study the binding of POT1 to physiological telomeric tails using AFM we utilized two DNA substrates, Tel16 and Ctrl16 (**Table 2.1**). Ctrl16 has the same length as the Tel16 DNA substrate, but every other TTAGGG sequence in Ctrl16 is changed to TTAGTG, which eliminates G4 folding (**Fig. 2.2E**)(123), but preserves POT1 binding(1). Accordingly, both Tel16 and Ctrl16 substrates have a maximum of eight POT1 DNA binding sites. Electrophoresis mobility shift assays (EMSA) showed that under the same conditions POT1 binds Tel16 and Ctrl16 substrates to a similar extent (**Fig. 2.9**). The appearance of more than one shifted band suggests that multiple POT1 molecules can bind to the Tel16 or Ctrl16 substrates. In the AFM images of Ctr16 with POT1, arrays of tandem POT1 proteins were observed (thin arrow, **Fig. 2.10A and C**), which were not present in the POT1 alone images (**Fig. 2.7A**). The mean height of these POT1 arrays is 0.83 (± 0.15) nm which is comparable to the POT1 height (0.71 ± 0.11 nm) in protein alone images. The statistically significant height difference ($p < 0.004$) between POT1 (0.83 ± 0.15 nm, on Ctrl16) and G4 DNA on Tel16 (1.32 ± 0 .22 nm), (compare **Figs. 2.3**,**2.6**,and 2.**7A,**) is due to the increased stiffness of G4 DNA compared to POT1 protein, which leads to less compression by the mechanical AFM imaging process. Accordingly, the distinct heights of POT1 and Tel16 in the AFM images provide a robust criterion to differentiate between POT1 and G4 structure when POT1 and Tel16 are mixed together (**Fig. 2.6B**). Previously, AFM height measurement has been used to differentiate a DNA repair protein UvrB and a UvrB-quantum dot conjugate(111)

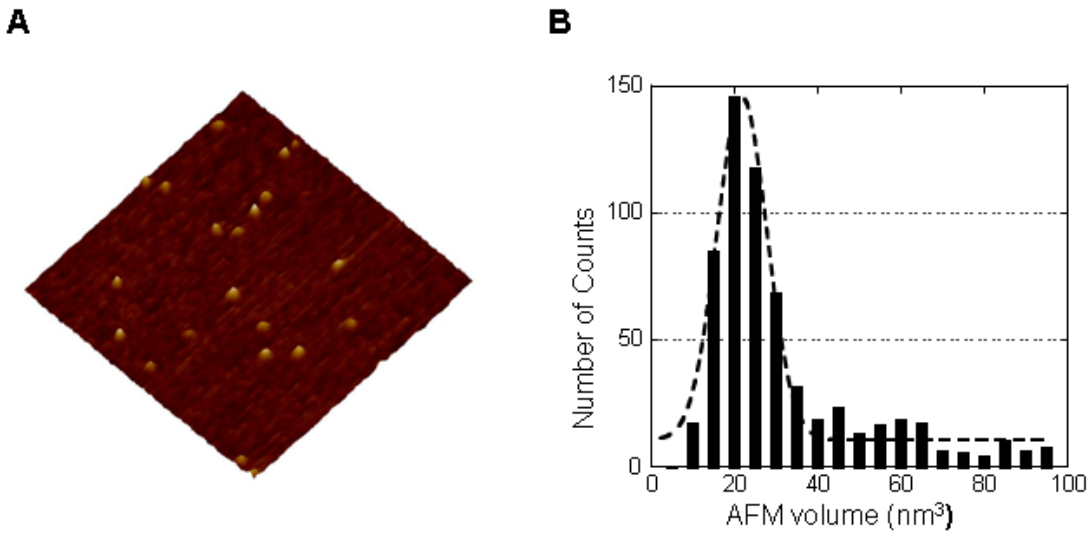

Fig. 2.7 Full length POT1 is a monomer in solution. (**A**) A representative AFM image of untagged POT1 protein at 200 nM concentration. The image is 1 μm × 1 μm at 2 nm height scale. (**B**) AFM volume distribution of POT1 from images of POT1 at 200 nM concentration. The dashed line represents the Gaussian fit to the data (n=664 molecules, R^2= 0.96), which is centered at 22 nm^3 and corresponds to POT1 monomer based on the standard calibration curve (**Fig. 2.14**).

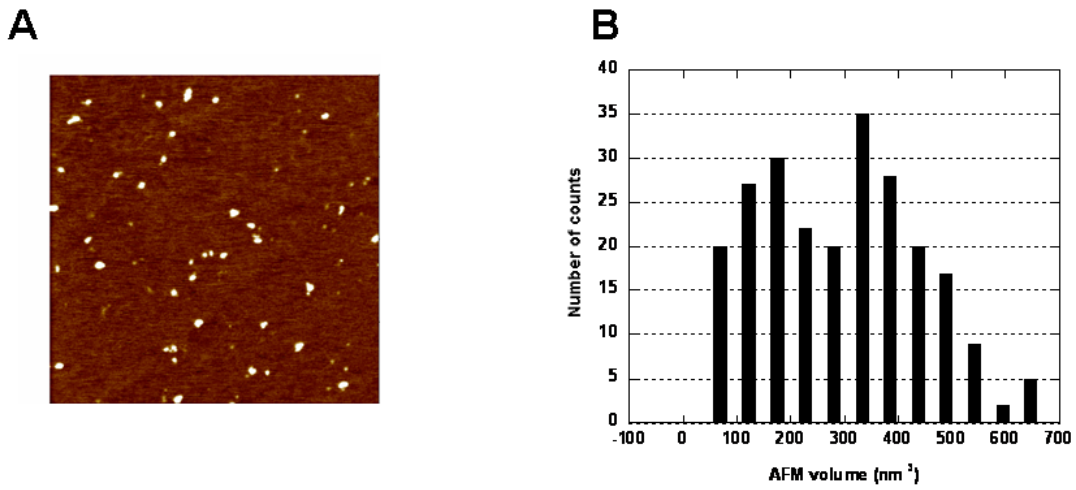

Fig. 2.8 AFM image and volume analysis of GST-POT1. (A) Representative AFM image of GST-POT1 protein at 20 nM concentration. The image size is at 1 μm × 1 μm. (B) AFM volume distribution of GST-POT1.

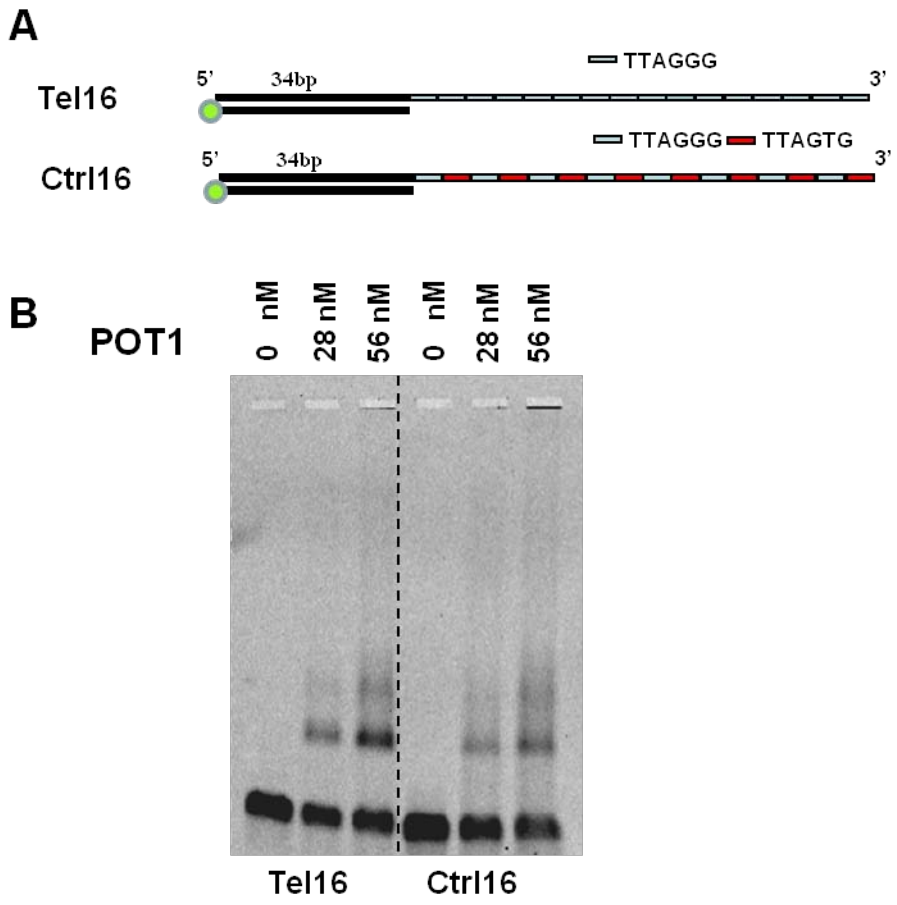

Fig. 2.9 Gel-shift assay of POT1 binding. (**A**) Schematic drawing of the DNA substrates for EMSA. The sequences of Tel16 and Ctrl16 are shown in Supplementary Table 1. The green circles represent 3' fluorescein adducted thymine. (**B**) EMSA assay of POT1 binding to Tel16 and Ctrl16 substrates. POT1 at indicated concentrations was incubated with Tel16 or Ctrl16 at 10 nM in 1X POT1 buffer at 37 ^{o}C for 15 minutes. Half of each reaction was removed immediately and loaded onto a 1% agarose gel and was subject to electrophoresis at 10 V/cm in 1X TBE buffer for 1 hour at 4 °C. The gel was scanned using a Typhoon^{TM} 9400 VariableMode Imager.

When 200 nM POT1 was incubated with Tel16 (1 μM) the percent of molecules that exhibited structures with G4 DNA character (height > 1 nm) was greatly reduced from 100% of the measured molecules of Tel16 alone to 24% (98/410) of the molecules visualized after co-incubating Tel16 with POT1 (**Fig. 2.10B**), and the majority of molecules (77%, 330/405) showed only structures that were characteristic of bound POT1. Importantly, of the G4 DNA structures observed (98/410), 23 molecules displayed multiple peaks with differing heights that were consistent with G4 DNA and bound POT1 on the same molecule (compare **Fig. 2.10C** with **D**). These images indicate that G4 DNA and POT1 can coexist on the same molecule. Moreover, the lengths of the G4 regions remaining after POT1 addition were less than half of the G4 region lengths on Tel16 alone (**Fig. 2.10E**), consistent with a POT1-mediated shift from two to one G4 units on each individual Tel16 molecule. It is also worth noting that the bound POT1 region length distributions for Ctrl16 and Tel16 (**Fig. 2.10F**) were very similar in that both exhibited a long right-sided "tail" representing similar numbers of POT1 proteins bound to Tel16 and Ctrl16 molecules. The length of longer POT1 arrays (45 to 60 nm) is consistent with the length of ssDNA (48 nm, assuming ssDNA as 0.5 nm/base) on fully extended Tel16 (**Fig. 2.10F**). Together, these data indicate that POT1 binding can successfully compete with G4 DNA folding on telomeric ssDNA.

Figure 2.10

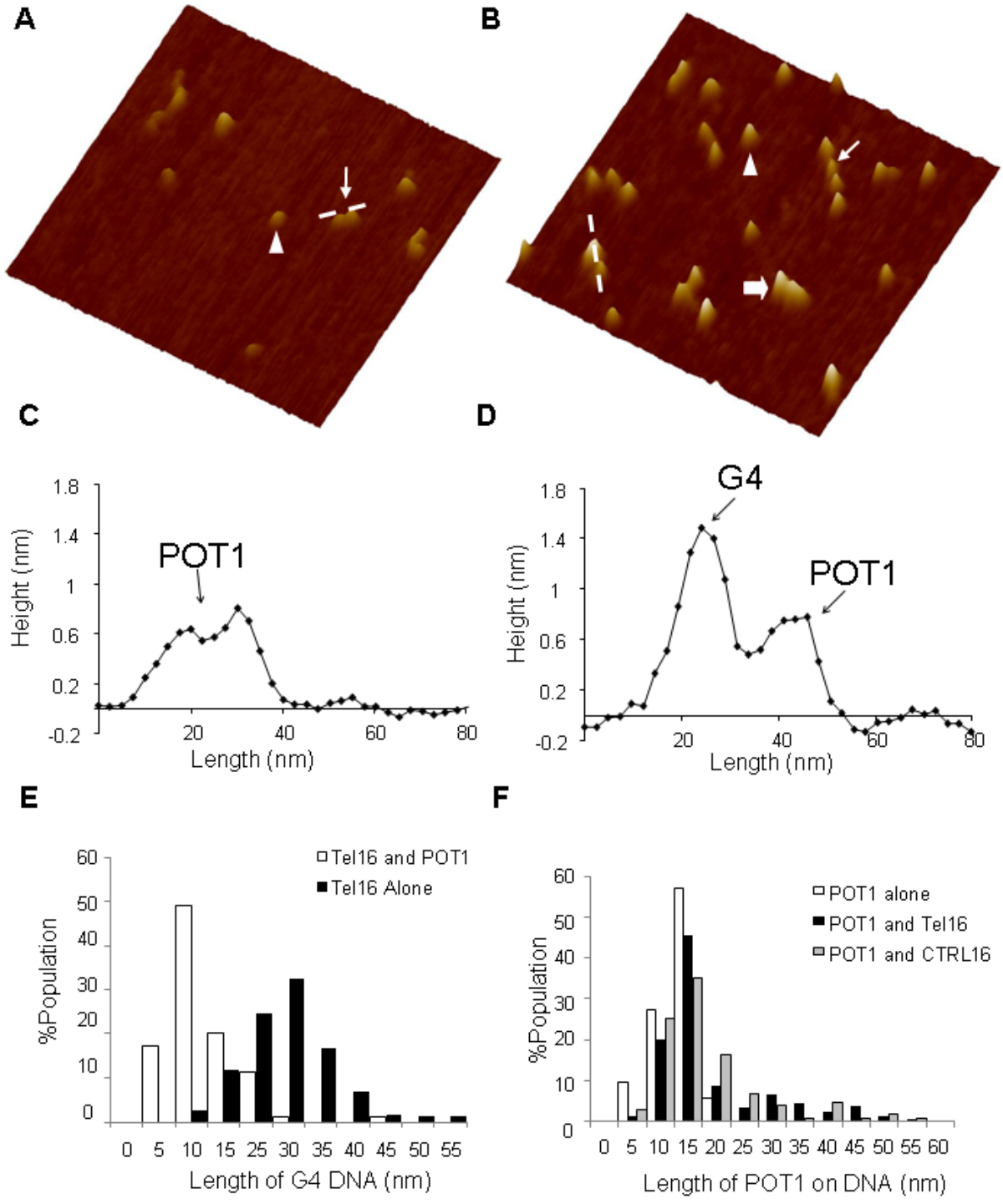
A
B
C
D
POT1
Height (nm)
Length (nm)
G4
POT1
E
F
Tel16 and POT1
Tel16 Alone
%Population
Length of G4 DNA (nm)
POT1 alone
POT1 and Tel16
POT1 and CTRL16
Length of POT1 on DNA (nm)

Fig. 2.10 POT1 competes with G4 DNA folding on 3' telomeric tails. (**A**) A representative AFM surface plot of Ctrl16 (1 μM) in the presence of POT1 (200 nM). (**B**) A representative AFM surface plot of Tel16 (1 μM) in the presence of POT1 (200 nM). The AFM images are 350 nm× 350 nm at 2 nm height scale. The triangles point to individual POT1 molecules. The thin arrows point to representative structures with multiple POT1 proteins. The thick arrow points to a representative structure with folded G4. Dotted lines denote molecules whose cross-sections are shown in **C** and **D**. (**C**) Cross section of molecule highlighted in **A** showing two POT1 molecules on the same telomeric tail. (**D**) Cross-section of molecule highlighted in **B** demonstrating that G4 (left) and POT1 (right) coexist on the same molecule. (**E)** Histogram of the length of G4 DNA regions (stretch of DNA with height taller than 1 nm) measured from AFM images of Tel16 alone (black bar) and Tel16 (1 μM) in the presence of POT1 (200 nM, open bar). (**F**) Histogram of the length of POT1 regions (stretch of structure feature with height between 0.5 and 1 nm) measured from AFM images of POT1 alone (200 nM, open bars; n=51 molecules), in the presence of Ctrl16 (1 μM, gray bars; n=103), and Tel16 (1 μM, black bars; n=330 molecules).

2.3 DISCUSSION

POT1 binding to $(TTAGGG)_4$ substrates prevents G4 DNA folding(103;124). However, the arrangement of G4 DNA and the competition with POT1 binding on a long, physiologically realistic telomeric substrate was unknown. In this study we used single-molecule imaging to examine the assembly of G4 units on DNA substrates containing 4 (Tel4), 8 (Tel8) and 16 (Tel16) TTAGGG repeats, with the latter representing the mid range of the telomeric overhang length in human cells(1). Telomeric DNA with well-defined lengths allowed us to study the length dependent formation of G4 structures at the single-molecule level. We demonstrated that G4 DNA assemblies on physiologically relevant telomeric tails rarely form the maximum potential number of G4 units, and G4s are arranged in a beads-on-a-string conformation. This study is the first to show via AFM imaging that POT1 is monomeric and stabilizes the ssDNA, driving the $(TTAGGG)_{16}$ structural equilibrium towards an extended, protein-bound state. Importantly, bound POT1 coexists in some cases with G4 DNA on the same Tel16 molecule. Together our data indicate that POT1 binding to unfolded segments of telomeric tails may destabilize existing G4 structures on that tail, thereby enabling the telomere t-loop/D-loop conformation or extension by telomerase.

The AFM imaging results from this study indicate that the majority of telomeric tails at lengths that occur in human cells did not fold into the maximum potential number of G4 units. The structure length and volume measurement comparison with Tel4, which can only form one G4 unit, indicates that the Tel8 and Tel16 molecules primarily fold into one and two G4 units, respectively (**Fig. 2.2D** and **2.8**). We applied a first principles combinatoric approach to understand the mechanism underlying the underfolding, and found that the prediction from this model for ssDNA with 13 repeats (**Fig. 2.4B**) is consistent with a bulk circular dichoism study of

a 13-repeat telomeric oligonucleotides in which only two quadruplexes formed on average(79). However, the normalized G4 distributions of Tel16 images demonstrated a sharp peak at two quadruplexes, whereas the probabilistic model based on the first-principles combinatoric approach predicted a nearly equal quantity of molecules with three quadruplexes as well (**Fig. 2.4B**). The discrepancy between our experimental observations and the probabilistic models may be explained by differences in the probability of forming G4 at different positions along the length of Tel16. A previous study using DMS footprinting and exonuclease hydrolysis with $T_{24}(TTAGGG)_7$ DNA substrates revealed that the probability of forming G4 rapidly decreases towards the 5' flanking sequence(98), from 55.8% at the 3' end (0 position) to 21.8%, 14.5% and 7.9% at the first, and second and third positions (next to 5' flanking sequence), respectively. Our model (**Fig. 2.4B**) presumes that probabilities of forming G4 along the 3' G-rich tail of Tel16 are the same. The dramatic decrease in the probability of forming G4 units when the repeat positions are close to the 5' flanking region, effectively shortens the number of available repeats for G4 folding on Tel16. This explanation is supported by the close agreement of the normalized G4 distributions from the experimental data with the theoretical G4 distributions of two shorter substrates with 13 and 14 repeats, repeats (**Fig. 2.4B**), the former of which was experimentally verified to preferentially form two quadruplexes(79). A previous report indicated that while $GGG(TTAGGG)_3$ is the most stable, as repeat number increases (n=7-16) the quadruplex molecules become less thermostable(77). The presence of loops with various length on the tetraplex sides can potentially lead to irregularities in G4 structure and, consequently, to the structure destabilization. Current literature suggests that loop length and composition strongly

Figure 2.11

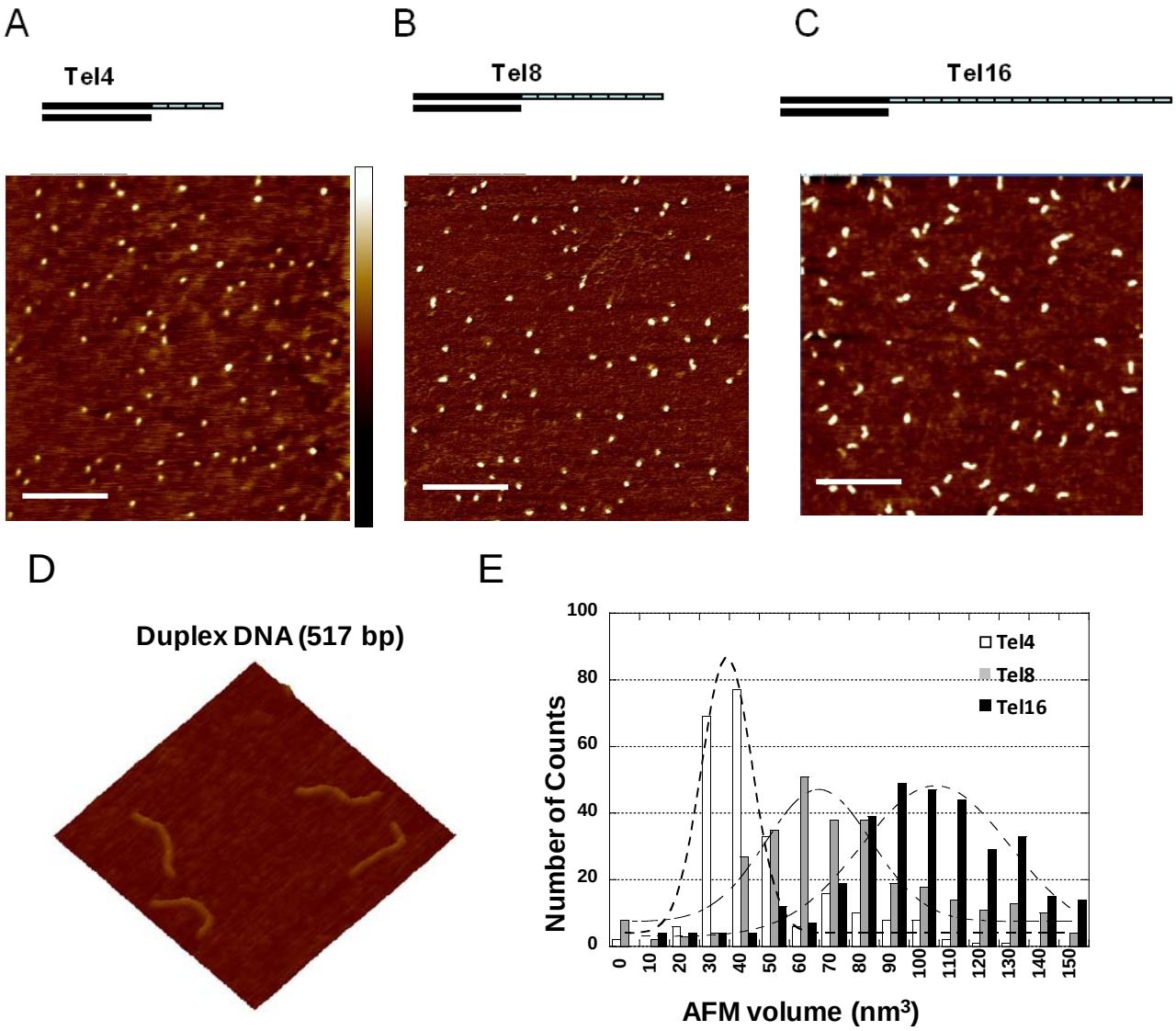
A
Tel4
B
Tel8
C
Tel16
D
Duplex DNA (517 bp)
E
Number of Counts
AFM volume (nm³)
Tel4
Tel8
Tel16
0
20
40
60
80
100
0
10
20
30
40
50
60
70
80
90
100
110
120
130
140
150

Fig. 2.11 Telomeric constructs, PCR fragments, and AFM volume analysis. (**A)** to (**C)** are representative field view images of Tel4, Tel8 and Tel16, respectively. The images are 1 μm × 1 μm. The color bar in A represents heights from 0-2 nm. The scale bars are 250 nm. (**D**) A representative AFM surface plot of 517 bp long PCR fragments. (**E**) The histogram of volume measured from AFM images of Tel4 (open bars), Tel8 (gray bars) and Tel16 (black bars). The dotted lines represent Gaussian fits to the data. The mean AFM volumes of Tel4, Tel8 and Tel16 are: 40 (± 2), 57 (± 4) and 106 (± 3) nm^3 (3 depositions for each substrate), respectively. It is worth noting that the mean volume of Tel16 is slightly larger than simple addition of two single G4 measured from Tel4 images. This is partly due to the inclusion of unfolded structures alongside with G4 on individual Tel16 molecules since AFM derived volume is a measurement of total volume above the noise level (see **Methods**).

influence the quadruplex stability, and quadruplexes formed by $(TTAGGG)_5$ with a 9-nt-loop loops were less stable than the one made of 4 consecutive repeats(98).

The arrangement of G4 DNA on longer physiological telomeric tails has been controversial. One melting study supported a beads-on-a-string conformation whereby long telomeric substrates fold into the maximum potential number of quadruplexes and the quadruplexes do not directly interact with each other(78). Another study found that a stacked model whereby individual quadruplexes fold in a way that their loop reactions interact, and a more rigid superstructure is formed(105;119). Direct visualization of individual molecules in our study revealed that 23% and 1% of the measured Tel16 molecules had two and three discernable peaks, respectively. These results support a beads-on-a-string model whereby the quadruplexes form as individual G4 units separated by stretches of ssDNA, creating a more flexible structure with discernable peaks (**Figs. 2.2**, **2.3, 2.12**). While not all the molecules displayed distinct peaks, this was likely due to the resolution limits of the AFM under the current imaging conditions. If two quadruplexes are linked by a TTA linker, the AFM cannot resolve two individual peaks; roughly 1.5 telomeric repeats are required to resolve two peaks (for the calculation of AFM resolution see **Chapter 4.1**). Also, while the average height of the Tel16 molecules was 1.32 (± 0.22) nm, the average height of the G-wires was 1.63 (± 0.17) nm, suggesting that Tel16 G4 DNA is more flexible, corroborating a beads-on-a-string arrangement. Previous studies indicated that POT1 binding to substrates with four repeats trapped the molecules in an extended state, shifting the equilibrium from a folded G4 unit to an extended conformation(103;124). However, the arrangement of G4 DNA and the competition with POT1 binding on a long, physiologically realistic telomeric substrates was unknown. Our finding that the majority of Tel16 molecules only form two G4 structures has important implications for

POT1 loading on physiologically relevant telomeric tails. POT1 can not bind the short $(TTAGGG)_4$ substrates until the equilibrium shifts from G4 structure to an extended state(103). In contrast, on the physiologically relevant Tel16 substrates, an underfolded Tel16 molecule constantly has multiple ssDNA sites available for POT1 binding; therefore, POT1 loading does not require thermal melting of existing G4 DNA.

We propose that POT1 binds to the unfolded ssDNA regions, and sterically impairs adjacent telomeric repeats from folding into G4 DNA, thereby promoting unfolding into extended ssDNA (**Fig. 2.13B**). This is in contrast to the previous passive model, in which telomeric tails can be maximally folded in G4 DNA and POT1 is incapable of binding until the equilibrium shifts from G4 to ssDNA (**Fig. 2.13 A**). The steric driver model is consistent with results from AFM imaging of Ctrl16 and Tel16 with POT1 (**Fig. 2.10**). Specifically, upon addition of POT1, the equilibrium shifts from a majority of Tel16 molecules forming two quadruplexes to one quadruplex and/or multiple POT1 monomers bound (**Fig. 2.10**). Importantly, multiple POT1 molecules bind Tel16 and the non-G4 forming Ctrl16 substrate to similar extents, leading to protein arrays of roughly equal length distributions (**Fig. 2.10F**). If POT1 can only capture the ssDNA when the equilibrium shifts from G4 DNA to ssDNA, then we would expect a greater number of molecules with long POT1-bound regions for Ctrl16 relative to Tel16, since POT1 does not need to compete with G4-folding to bind Ctrl16.

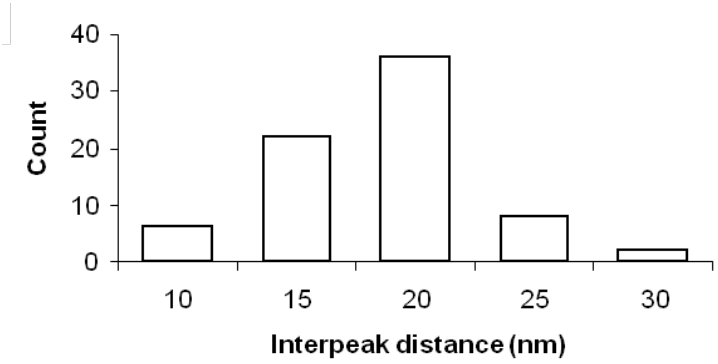

Fig. 2.12 Interpeak distance distributions on Tel16. Examples of the measurement of interpeak distance are shown in Fig. 3. The mean interpeak distance, 20 nm, corresponds to ~40 nucleotides (~7 telomeric repeats, assuming 0.5 nm/per base).

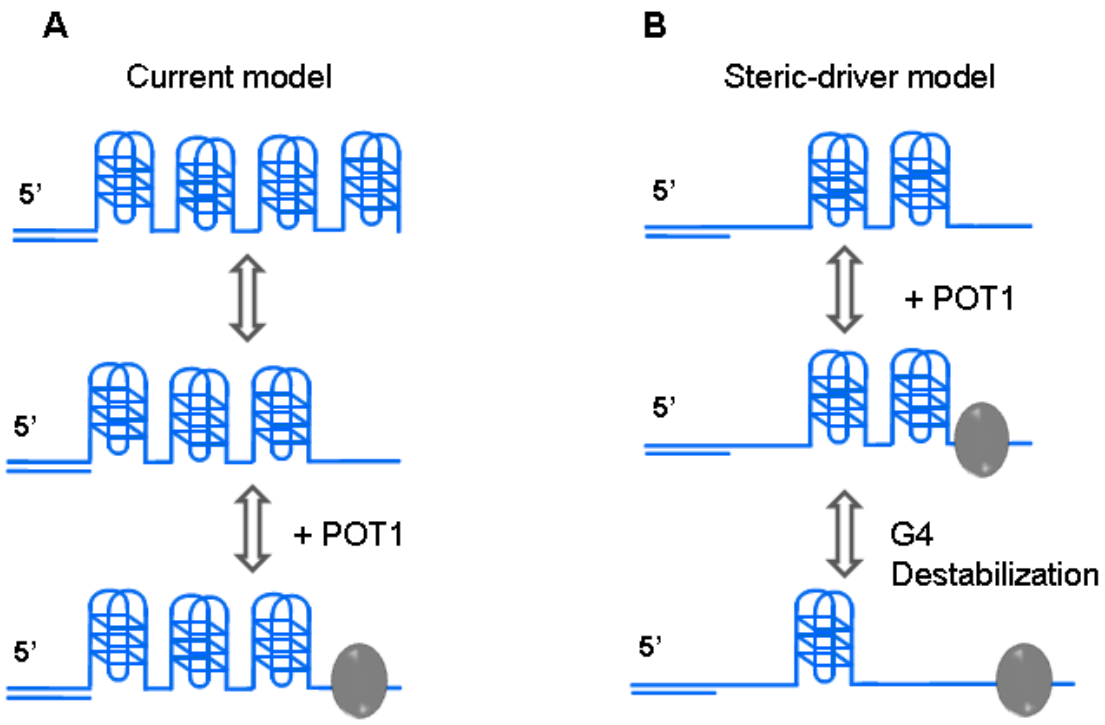

Fig. 2.13 Comparison of the passive and steric driver models of POT1 binding. A 3' telomeric tail with 16 TTAGGG repeats is drawn as an example. POT1 is shown as gray ovals bound close to 3' end. (**A)** A passive model whereby G4 DNA is arranged on a beads-on-a-string, maximally-folded telomeric tail. POT1 could not bind the telomeric sequence until the equilibrium shifts from G4 to an extended state. Then POT1 traps the molecule in the extended state. (**B**) In a steric driver model, at equilibrium the telomeric tails rarely form the maximum potential number of G4 units. POT1 is able to bind unfolded telomeric repeats and destabilize existing G4 DNA on the same molecule. Multiple POT1 molecules can bind to the different stretches of ssDNA on the same 3' telomeric tail and impose an emergent synergistic effect against G4 folding.

In summary, we propose a model whereby POT1 acts not as an active DNA unwinder, but rather as a "steric driver" by binding to underfolded telomeric tails and thereby destabilizing the remaining G4 structures on the molecule (**Fig.2.13B**), as evidenced by the reduction of G4 DNA structures upon POT1 addition (**Fig. 2.10E**). Our results provide a mechanism whereby a passive DNA binding protein may still impose an emergent synergistic effect against G4 DNA folding, which is consistent with previous studies on *S. pombe* POT1 DNA-binding studies(125;126). POT1 binding competition with G4 DNA folding on physiologically relevant 3' telomeric tails suggests an important mechanism for preserving telomere stability. At the late G_2 phase, POT1 levels at the telomere ends decrease, and the telomeres are temporarily recognized as DNA damage and are unprotected by the telomere complex before POT1 relocalizes to the telomeres(127). Since the unprotected tail can spontaneously fold into G4 DNA, this raises the issue of how POT1 reloads on the telomeric tail(128-130). Another study demonstrated that a G4-stabilizing agent induced an ATR-dependent DNA damage response, but that POT1 levels at the telomere ends remained unchanged(131), implying that G4 DNA and POT1 may coexist at telomere ends. The AFM images in this study show that the underfolding (i.e. formation of less than the maximum potential number of G4 units) of long telomeric ssDNA provides a route for POT1 binding and that POT1 and G4DNA can coexist on the same 3' telomeric tail. The direct visualization of single molecules that resemble physiologically relevant telomeric tails provide a mechanistic basis for understanding the modulation of telomere structure and function by POT1 and G4 DNA.

2.4 MATERIALS AND METHODS

DNA Substrates. All oligonucleotides were purchased from Integrated DNA Technologies and were purified using PAGE by the manufacturer. The sequences of the oligos are listed in Supplementary Table 1. DNA substrates which contain a 5' duplex region and a 3' single-stranded tail were formed by incubating equal molar amounts of oligonucleotides in 1X phosphate buffer (10 mM potassium phosphate and 150 mM KCl) or 1X POT1 buffer (40 mM Hepes pH 7.5 and 50 mM KCl) at 85^{o} C for 5 mins, followed by slow cooling to room temperature.

Protein p urification. Recombinant GST tagged and untagged POT1 proteins were purified using a baculovirus/insect cell expression system and an AKTA Explorer FPLC (GE Health Care, Piscataway, NJ) as described previously(132). Protein concentrations were determined using Coomassie staining along with a standard of known concentration. Proteins used in this study are more than 90% pure based on SDS-PAGE and Coomassie staining.

AFM sample preparation and imaging. All DNA substrates and POT1 protein were diluted in 1X POT1 buffer containing additional 10 mM $MgCl_2$ for AFM imaging. All buffers were heated at 65 degree for 15 to 30 mins to dissolve small salt particles accumulated during storage. When POT1 was present, DNA and proteins were incubated at 37^{o} C for 10 mins. G-wire solution was prepared by incubating a 270 μM solution of $G_4T_2G_4$ monomer in 100 mM potassium phosphate buffer (pH 7) at 90 ^{o}C for 10 minutes and slow cooling to room temperature, followed by incubation at 4 ^{o}C for 12 hrs. All samples for AFM imaging were prepared by depositing samples onto a freshly cleaved mica (SPI Supply, West Chester, PA), followed by washing with MilliQ water and drying under a stream of nitrogen gas. All images were collected using a MultiMode V

microscope (Veeco Instruments, Plainview, NY) using Escanners in tapping mode. Pointprobe® plus noncontact/tapping mode silicon probes (PPP-NCL, Agilent) with spring constants of ~50 N/m and resonance frequencies of ~190 kHz were used. Images were captured at a scan size of 1 μm × 1 μm, a scan rate of 2-3 Hz, a target amplitude of 0.3 V and a resolution of 512 × 512 pixels.

The combinatoric model for G4 formation. Statistical analysis of G4 formation on Tel8, Tel13, Tel14, Tel15 and Tel16 were calculated by treating them as a sequence of 8 and 13-16 lattices, respectively. It was assumed that G4 structures can form by four consecutive TTAGGG repeats and that individual G4s can fold randomly along the entire length of the lattice. The number of possible arrangements of the *h* items (G4s and unstructured repeats) can be described as: ${}^{h}C_{i} = \frac{h!}{i!(h-i)!}$ where *i* is the number of G quadruplexes.

E.g., for Tel8, there are ${}^{5}C_{1} = \frac{5!}{1!(5-1)!} = 5$ ways to arrange a single quadruplex, and only one way to arrange two quadruplexes.

Statistical analysis of AFM images. Evaluation of the AFM height at different target amplitudes indicated that within the range of target amplitudes used in this study (300 to 400 mV), the height variation in AFM images of an antibody which is similar in scale to the protein and DNA molecules in this study, is ~15% of the total height (**Supplementary Fig. S3A**). The length measurement was done using the Nanoscope7.30 software; structures over 1 nm were noted as G4 DNA, whereas POT1 peaks fell between 0.5 and 1 nm. Two discernable G4 peaks on Tel16 were defined as presence of two local maxima over 1 nm with a trough in between which was at least 0.2 nm lower than the shorter peak (**Fig. 3B**). For AFM volume analysis, dimension of proteins were measured using Image SXM software(110;116;122). AFM volume of

a particle was calculated as $V = S \times (H - B)$, where V is the AFM volume, S is the area generated at the base of a protein using “density slice” function of the SXM software, H is the average height, and B is the background height.

Name	Sequence (5’ to 3’)
Tel4-top	**A GGT ATC AGC ATA ATG GCC ACG GTG CGT ACT GCG** $(TTAG_3)_4$
Tel8-top	**A GGT ATC AGC ATA ATG GCC ACG GTG CGT ACT GCG** $(TTAG_3)_8$
Tel16-top	**A GGT ATC AGC ATA ATG GCC ACG GTG CGT ACT GCG** $(TTAG_3)_{16}$
Ctrl16-top	**A G GT A TC A GC ATA AT G GCC AC G GTG C GT ACT GCG** $(TTAG_3TTAGTG)_8$
Bottom	**CGC AGT ACG CAC CGT GGC CAT TAT GCT GAT ACC T**
G-wire	G_4TTG_4

Note: Complementary regions are shown in bold.

Duplex DNA substrates are annealed using the following pairs of oligos:

Tel4=Tel4-top/Bottom; Tel8=Tel8-top/Bottom; Tel16=Tel16-top/Bottom;

Ctrl16=Ctrl-top/Bottom.

Table 2.1 Oligonucleotides used in Chapter 2.

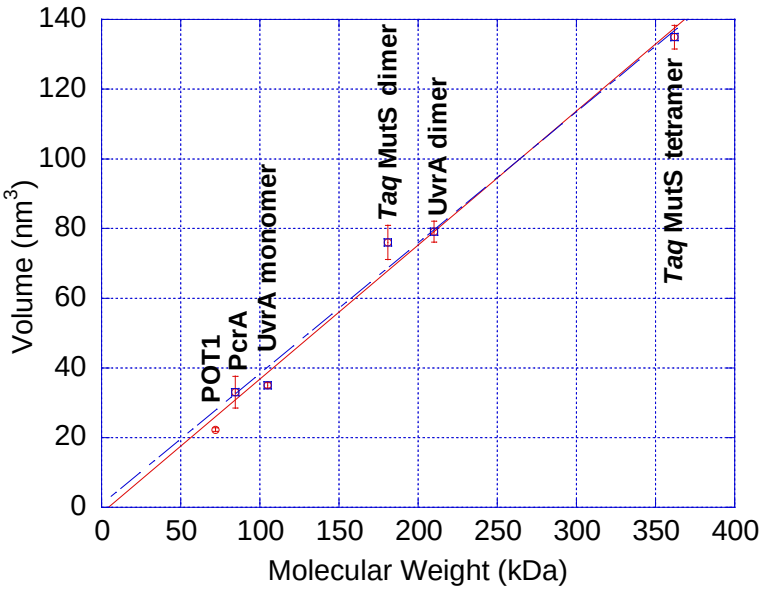

Fig. 2.14 The AFM volume of a globular protein varies linearly with its molecular weight. The volumes were determined as described previously(122). The plot corresponds to data for the following proteins and their various oligomeric states: PcrA monomer (86.4 kDa), UvrA monomer (105 kDa), Taq MutS dimer (181 kDa), UvrA dimer (210 kDa), and Taq MutS tetramer (362 kDa). The error bars represent the standard deviation from 2-3 depositions for each protein. The blue line corresponds to the least-square fit of the data not including POT1, which is described by the following equation: V (nm^3) = 0.3735 MW (kDa) + 1.435 (R^2=0.9945). Based on this equation, the predicted AFM volume of POT1 (72 KDa) is 28 nm^3. The red line corresponds to the fit of the data including POT1 (R^2=0.9883).

3.0 TELOMERIC PROTEIN TRF2 PROTECTS HOLLIDAY JUNCTIONS WITH TELOMERIC ARMS FROM DISPLACEMENT BY THE WERNER SYNDROME HELICASE

The following chapter has been published by *Nucleic Acids Research* (2010, doi: 10.1093/nar/gkq144) with authors Gerald J. Nora, Noah A. Buncher, and Patricia L. Opresko.

.

3.1 INTRODUCTION

Lack of the Werner protein (WRN) causes Werner Syndrome (WS), a segmental progeroid disorder characterized by cellular genomic and telomeric instability and premature senescence (7;11). WRN is a RecQ family helicase that is unique among the five human RecQ helicases for also having an exonuclease domain (20). Many of the RecQ helicases in general are thought to function during DNA replication to prevent replication fork demise and to restore stalled or broken replication forks, partly through homologous recombination (HR) pathways (133). Consistent with this, WRN protein is implicated in pathways for recombinational repair of stalled replication forks and DNA double strand breaks (44;134). WRN is proposed to have particularly important roles in telomere preservation during replication. Telomeres protect chromosomes ends, and telomere dysfunction triggers cellular senescence, apoptosis, or genomic instability (1). The forced expression of telomerase in WS fibroblasts suppresses many of the primary cellular defects including telomere loss on sister chromatids (18), premature senescence (17), and the accumulation of chromosomal aberrations (135). Cellular and biochemical data indicate that WRN may preserve telomeres by dissociating alternate DNA structures to facilitate replication fork progression or for completion of HR repair at stalled or broken replication forks (reviewed in (7)). These DNA structures include G-quadruplexes, 4-way junctions that mimic regressed replication forks and Holliday Junction (HJ) intermediates in HR, and D-loop structures that occur during HR and at the telomeric end (30;34). All of these are preferred substrates for WRN helicase which can also branch migrate HJ and D-loops (34;136). These studies suggest that inappropriate processing of alternate structures at telomeres can lead to premature telomere loss and cell senescence or apoptosis.

Human telomeres are characterized by repetitions of a short sequence of duplex DNA (TTAGGG) and a 3' overhang of single stranded DNA (ssDNA), bound by the shelterin complex of six proteins (1;137). The 3' overhang, which may be 50-500 nt, is bent around into a t-loop and then invades the duplex DNA, displacing the G-rich strand and forming a displacement loop (D-loop) (38). Telomere D-loop formation is thought to be mechanistically similar to D-loop formation during the initiation of homologous recombination (1;36;37). The shelterin protein TRF2 is critically important for stimulating formation and preservation of the telomeric t-loop/D-loop structure (57;138;139). TRF2 has two DNA binding domains: an N-terminal basic domain that binds four way junctions regardless of sequence (56) and a C-terminal Myb domain that specifically binds duplex TTAGGG repeats (55). Cellular studies showed that the overexpression of TRF2 lacking the B-domain (TRF2ΔB) results in cleavage of the t-loop/D-loop by HJ resolvase XRCC3 (37), which leads to extra-chromosomal telomeric circles that can be visualized by 2-D electrophoresis and electron microscopy (37;140). TRF2 binds to and inhibits cleavage of 4-way junction models of HJ structures *in vitro* (56;84). WRN is required for the production of telomeric circles in TRF2ΔB overexpressed cells (83), but the mechanism is unknown. Interestingly, in telomerase positive WS cells, telomeric circles form in the absence of WRN and independently of XRCC3, implicating WRN in more than one pathway for telomere stability (83). While TRF2 is known to protect HJs from cleavage by HJ resolvases (84), whether TRF2 also regulates displacement and migration at the telomere t-loop/D-loop is not known.

One model for WRN promotion of telomeric circles in TRF2ΔB expressing cells is that WRN stimulates branch migration of the t-loop/D-loop into a target for cleavage. HJ-like structures also form preferentially at telomeres during replication fork regression that produces a

four-way junction in which one duplex end is accessible (chicken foot) (80). WRN exonuclease can attack this vulnerable end, and the helicase can displace the 4-way junction (141). Both TRF2 and WRN bind the HJ core *in vitro* (56;142). Therefore, it is possible that TRF2 may protect HJs from WRN activity by inhibiting WRN loading. Paradoxically, TRF2 binds directly to WRN protein *in vivo* (28;83;143), and *in vitro* studies showed that TRF2 recruits WRN to telomeric substrates (143), stimulates WRN exonuclease digestion of telomeric duplexes (143), and stimulates WRN helicase unwinding of short forks (28). Whether TRF2 modulates WRN processing of HJ DNA is not known.

Given that TRF2 represses WRN promotion of t-circles (83), in this study, we tested the hypothesis that TRF2 negatively regulates WRN activity on 4-way structures that may occur at the t-loop/D-loop or during replication fork regression. By testing a variety of HJ substrates containing site specific blocks to WRN helicase 3' to 5' translocation, we provide evidence that WRN displaces 4-way structures from the center moving outward. We also found that TRF2 protects telomeric HJ DNA against WRN strand displacement activity. However, both the TRF2 B and Myb DNA-binding domains are required for WRN inhibition, indicating that TRF2 must engage both the HJ core and the duplex arms to regulate WRN strand displacement.

3.2 MATERIALS AND METHODS

3.2.1 Proteins

Recombinant human hexahistidine-tagged WRN protein and the exonuclease-dead E84A WRN mutant (X-WRN) were purified from a baculovirus/insect cell expression system as described previously (144). Recombinant human hexahistidine-tagged RPA was provided as a gift from Dr. Walter Chazin (Vanderbilt University, TN, USA). Recombinant human hexahistine-tagged

TRF2 protein was purified as formerly described (28). Recombinant hexahistine TRF2 protein fragment (amino acids 45 to 501; TRF2ΔB) was generated by PCR cloning using the TRF2 cDNA from the baculovirus DNA construct, kindly provided by Dr. Titia de Lange (Rockefeller University, New York, NY) as a template. The PCR product was subcloned into the BamHI and EcoRI sites of the pRSET-A expression vector (Invitrogen, Carlsbad, CA). The TRF2ΔB fragment was expressed in BL21(DE3*) E. coli* and purified with an AKTA Explorer FPLC (GE Life Sciences, Piscataway, NJ). After the induction of TRF2ΔB expression with 1 mM IPTG for 4 hours at 30º C, cells were harvested and resuspended in lysis buffer (20 mM NaH_2PO_4, 0.5 M NaCl, 10 mM imidazole, 1% NP-40 and 5 mM beta-mercaptoethanol) and mixed on a rotator at 4ºC for 30 min. Protease inhibitors were included in all buffers (Roche Molecular Biochemicals, Indianapolis, IN). Cells were centrifuged at 15,000 rpm for 30 min at 4ºC. The supernatant was loaded onto a HisTrap FF (GE Life Sciences) column equilibrated with lysis buffer. The column was washed with Buffer A (20 mM NaH_2PO_4, 0.5 M NaCl) containing 10 mM imidazole, and subsequently washed and protein eluted with 60 mM and 100 mM imidazole, respectively. The eluant was dialyzed against buffer D (20 mM Hepes, pH 7.9, 100 mM KCl, 3 mM $MgCl_2$, 1 mM DTT, 20% glycerol, 0.5 mM PMSF) in Slide-a-lyzer cassettes (Thermo Fisher, Rockford, IL, USA) at 4º C. Protein purity and concentration were determined by SDS PAGE Coomassie analysis and Bradford assay, respectively. The pGEX4T-RAP1 construct was kindly provided by Dr. Zhou Songyang (Baylor College of Medicine, Houston, TX). GST-tagged human RAP1 was expressed in BL21 (DE3) *E. coli* and purified on glutathione Sepharose beads (GE Life Sciences) in batch followed by thrombin (GE Healthcare) cleavage as previously described with some modification (58). Following the induction of RAP1 expression with 1 mM IPTG for 2 hours at 37 ºC, cells were harvested at 3,800 rpm for 15 min at 4 ºC, resuspended in lysis buffer

(1X PBS, protease inhibitors, 50 mM beta mercaptoethanol, and 1% Triton X-100), sonicated, and mixed in a rotator for 15 min at 4ºC. Cell lysate was harvested at 10,000 rpm for 10 min, and the supernatant was incubated with glutathione-sepharose beads (GE Healthcare, Piscataway, NJ) for 3 hours at 4ºC on a rotator. Bound GST-tagged RAP1 was cleaved with Thrombin (GE Healthcare) for 12 hours at 4ºC. The supernatant was loaded onto a Mono Q 5/50 GL (GE Life Sciences) column equilibrated with Buffer A (150 mM Tris, pH 8.0, 50 mM NaCl, 0.05% NP-40, and 10% glycerol). RAP1 protein was eluted in Buffer B (150 mM Tris pH 8.0,10% glycerol, and 0.05% NP-40) at 335 mM NaCl. Protein purity and concentration were determined by SDS PAGE Coomassie analysis and Bradford assay, respectively.

3.2.2 DNA Substrates

All oligonucleotides were purchased from Integrated DNA Technologies (Coralville, IA) and PAGE-purified by the manufacturer, except for HJbio-4, which was purchased from and PAGE-purified by Gene Link (Hawthorn, NY, USA). Fluorescent labels were covalently linked to the 5' nucleotide by the manufacturer. Sequences are provided in Supplemental Table 1.

All HJ constructs, modified from the XJ12 described previously (30) and with the exception of HJT (Table 1), were annealed in 25 µL reaction volumes. For HJ DNA constructs, 10 or 12 pmol of the indicated oligonucleotide 1 was mixed with a 1:2:3:2 molar ratio with the respective 2, 3, and 4 oligonucleotides (Table 1), respectively, in 0.35X PBS and 350 mM LiCl. The oligonucleotides were annealed in a TC-312 PCR machine (Techne, Burlington NJ) by incubating at 95º C for 5 min, followed by cooling to 65º C at 1 deg./min. The reactions were held at 65º C for 5 min, cooled to 37º C at 1 deg./min, then incubated at 37º C for 8 h and then cooled to 10º C.

The HJT construct was prepared by first annealing the forked halves of the substrate separately in 12.5 µL reactions with 100 mM LiCl and 0.5X TE buffer: one with 10 pmol T-1 and 20 pmol T-2, and the other with 40 pmol T-3 and 20 pmol T-4. The reactions were incubated at 65º C for 5 min and allowed to cool to 37º C. To form HJT the separate forked duplex reactions were combined, incubated at 37º C for 1 h, and cooled to room temperature.

A portion of each HJ preparation (0.5 µL) was removed for use as a standard to determine the concentration of the purified HJ preparations. The remainder was mixed with 16 µL H_2O and 7.5 µL 80% glycerol and loaded on an 8% polyacrylamide gel. Voltage (150 V) was applied for 1.5 hours at 4º C. The HJ products were visualized by fluorescence on a Typhoon Imager (GE Healthcare), excised from the gel and purified from the gel slice by electroelution for two hours in a Model 422 Electroeluter (Bio-Rad, Hercules CA). The eluant was concentrated in a Microcon YM-30 (Millipore, Billerica MA) and resuspended in 100 µL of storage buffer (10 mM Tris pH 7.5, 10 mM $MgCl_2$) and stored at -20º C. Purification quality and yields were determined by analysis on 8% native polyacrylamide gels, followed by visualization and quantitation with a Typhoon Imager and ImageQuant software (GE Healthcare, Piscataway, NJ). Forked duplexes were prepared by annealing oligonucleotides as described in Supplemental Table 1 in 1:1 molar ratios. Annealing reactions (50 µL) were incubated in 50 mM $LiCl_2$ and 0.71X PBS at 95° C for five min and then allowed to cool to room temperature.

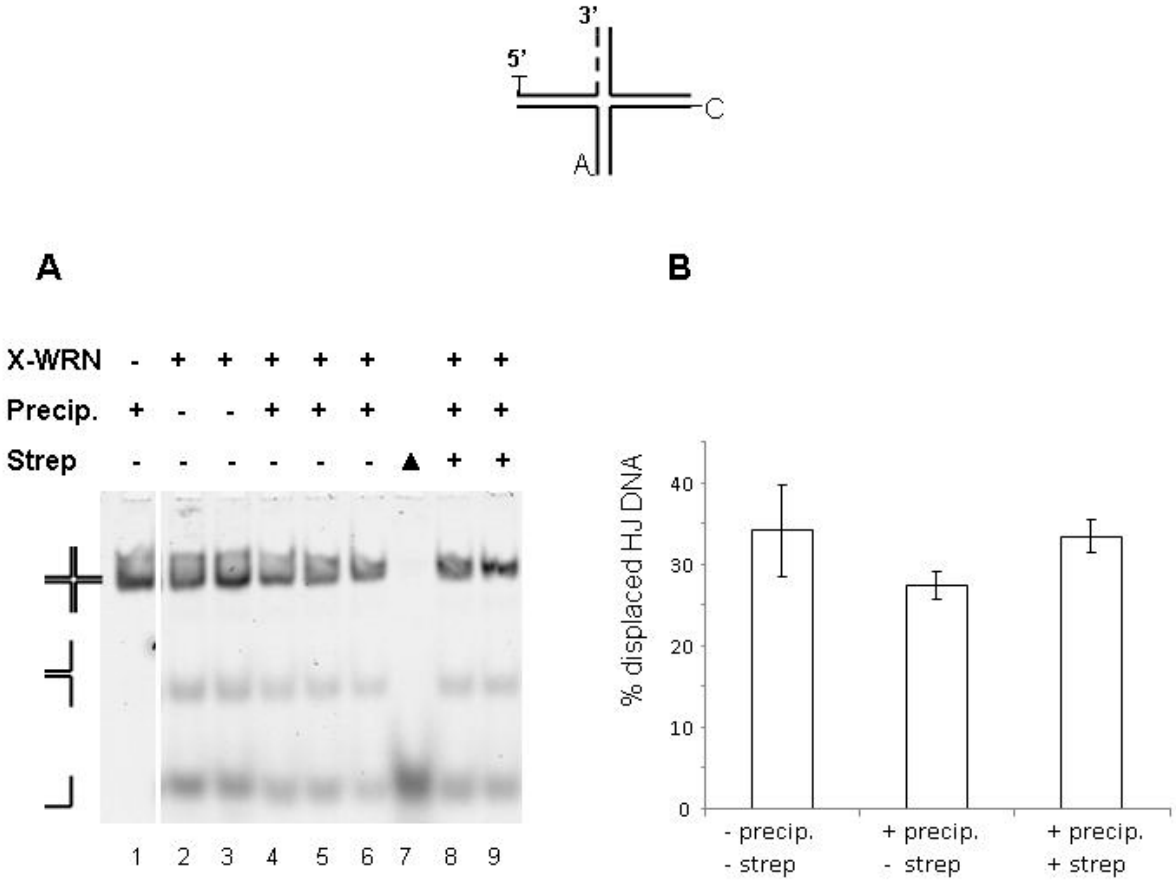

Fig. 3.1. SDS precipitation and streptavidin addition do not significantly alter WRN activity. Schematic of the HJA construct. T = TAMRA, C = Cy5, and A = Alexa488 labeled oligonucleotides. The 3' end of the exonuclease vulnerable T-oligo is highlight as a dotted line. (**A**) The HJA substrate (0.5 nM) was incubated under standard reaction conditions with 15 nM X-WRN (exonuclease-dead WRN) for 1 h. Streptavidin (30 nM) was included in the reactions for lanes 8 and 9. The reactions were terminated with 2.7x control buffer (lanes 2-3; identical to ProtK-SDS buffer only without LiCl; see **Materials and Methods**) or ProtK-SDS buffer and precipitated as described in **Materials and Methods** (lanes 5-9). ▲, boiled substrate. The reactions were separated on an 8% native gel and imaged in the Cy5 emission channel. (**B**)

Quantitation of the displaced HJ DNA. The percent of displaced HJ DNA was quantitated as described in **Materials and Methods**. Values represent the mean and standard deviation from two or three independent experiments.

3.2.3 Helicase/Exonuclease Reactions

HJ substrates were reacted in standard reaction buffer (40 mM Tris pH 7.5, 4 mM $MgCl_2$, 5 mM dithiothreitol (DTT), 0.1 mg/mL bovine serum albumin (BSA) and 2 mM ATP) at 37^{o}C for various time points as indicated in the Figure legends. The concentrations of substrate and proteins WRN, X-WRN, TRF2, TRF2ΔB and RPA were as indicated in the Figure legends. Reactions were started by WRN or X-WRN addition, and terminated by adding 2.7X stop dye (50 mM EDTA , 40% glycerol). The reactions that included RPA and/or TRF2 were terminated with 2.7X ProtK-SDS stop dye (15 mM EDTA, 1 µg/mL proteinase K, 300 mM LiCl, 1.5% SDS) and incubated at 37º C for 30 min. Because SDS fluoresces and interferes with analysis of the fluorescent substrates, it was precipitated based on a protocol by Zhitkovich and Costa (145). Briefly, 4 µL 1 M KCl was added to the 31.5 µl stopped reactions, followed by incubation on ice for 5 min. The precipitate was pelleted by centrifugation for 10 min at 10,000 x g at 4º C. The supernatant (31 µL) was drawn off and transferred to a fresh tube with 15 µL of loading buffer (40% glycerol, 0.5X TE). The SDS precipitation procedure did not result in any significant reannealing of the dissociated ssDNA products (**Fig. 3.1**). The terminated reactions were run on either 8 or 12% native polyacrylamide gels as indicated in the Fig. legends at 150 V.

The TAMRA, Cy5, and Alexa488 fluorophores were visualized on a Typhoon Imager using the preset TAMRA, Cy5, and Alexa488 laser excitation and emission settings, with a photomultiplier gain of 575 V for all channels and normal sensitivity. Control reactions ensured that cross talk between channels did not occur (**Fig. 3.4**). Reaction substrates and products were quantified using ImageQuant 5.1 (GE Life Sciences) software. The local median analysis was used for background correction. Substrate and products were calculated as the percent of the

total DNA in each lane for each fluorophore. All values were corrected for background in the no protein and boiled control lanes as described previously (33).

3.2.4 Electrophoretic mobility shift assay

Binding reactions (10 µl) contained standard reaction buffer (40 mM Tris pH 7.5, 4 mM $MgCl_2$, 5 mM DTT, 0.1 mg/mL BSA, 2 mM ATP), 5% glycerol, and 5 nM DNA substrate. The reactions were initiated by adding various concentrations of TRF2, as indicated in the Fig. legend, and were incubated for 20 min at room temperature. The reactions were loaded on 1% agarose gels and electrophoresed in 0.5X TBE at 4º C and 140 V for one hour. The fluorescent products and substrates were visualized and analyzed as described for the helicase reactions.

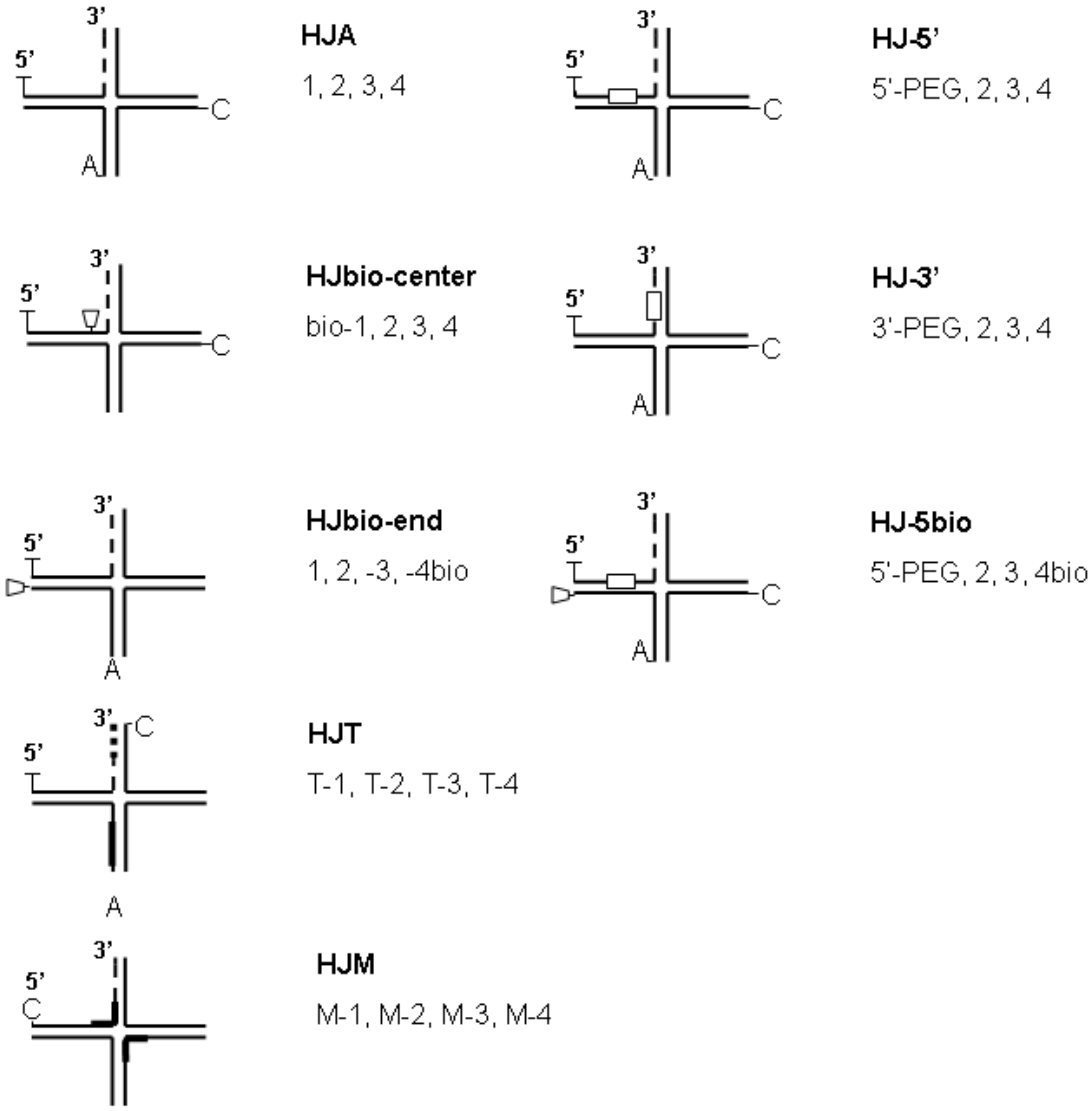

Table 3.1 Holliday Junction Substrates. T = TAMRA; C = Cy5; A = Alexa488. Dotted lines denote an oligonucleotide with an exonuclease-vulnerable 3' end lacking a thiophosphate bond protection. Bold lines indicate the G-rich strand of telomeric repeats. Trapezoid denotes biotin; open rectangle denotes PEG linker. Numbers denote oligonucleotides used (**Table 3.2**).

Table 3.2 List of oligonucleotides used in Chapter 3.

1	5'TMR-TACGCTGCCGAATTCTGGCTTGCTAGGACATCTTTGCCCACG TTGACCCA
2	5'TGGGTCAACGTGGGCAAAGATGTCCTAGCAATGTAATCGTCTATGACGT*C
3	5'Cy5GACGTCATAGACGATTACATTGCTAGGACATGCTGTCTAGAGACTATCG*A
4	5'TCGATAGTCTCTAGACAGGATGTCCTAGCAAGCCAGAATTCGGCAGCGT*A
	Biotinylated substrates
bio-1	5'TMR-TACGCTGCCGAATTCTGGCTTGC**Tbio** AGGACATCTTTGCCCACG TTGACCCA
3b	5'GACGTCATAGACGATTACATTGCTAGGACATGCAGTCTAGAGACTATCG*A
3bio	5'GACGTCATAGACGATTACATTGCTAGGACATGCAGTCTAGAGACTATCG*A-**bio**
4bio	5'AlexaTCGATAGTCTCTAGAC**Tbio**GCATGTCCTAGCAAGCCAGAATTCGGCAGCGT*A
	PEGylated substrates
PEG-3’	5'TMR-TACGCTGCCGAATTCTGGCTTGCTAGGACATCTTT-PEG-CACG TTGACCCA
PEG-5’	5'TMR-TACGCTGCCGAA-PEG-TGGCTTGCTAGGACATCTTTGCCCACG TTGACCCA
	Telomeric-repeat-containing (“hybrid”) substrates
C-1	5'TMR-TACGCTGCCGAATTCTGGCTTGCTAGGACAT**TTAGGGTTAGGGTTA GGG**
C-2	5'Cy5-**CCCTAACCCTAACCCTAAA**TG TCCTAGCAA**CCCTAACCCTAACCCTA*A**-3
C-3	5'Alexa-**TTAGGGTTAGGGTTAGGG**TTGCTAGGACATGCTGTCTAGAGACTATCG*A
C-4	5'TCGATAGTCTCTAGACAGCATGTCCTAGCAAGCCAGAATTCGGCAGCGT*A
T-1	5'TMR-TACGCTGCCGAATTCTGGCTTGCTGGACA**TTTAGGGTTAGGGTTAGG*G**
T-2	5'Cy5-**CCCTAACCCTAACCCTAAA**TGTCCTAGCAATGTAATCGTCTATGACGT*C
T-3	5'GACGTCATAGACGATTACATTGCTAGGACAT**CCCTA ACC CTA ACC CTA*A**-3'
T-4	5'Alex-**TTAGGGTTAGGGTTAGGGA**TGTCCAGCAAGCCAGAATTCGG AGCGT*A
M-1	5'Cy5-AATCATCGTCCTAGCAAGG**TTAGGGTTAGGG**GGCTGCTACCGGCACTGCG
M-2	5’CGCAGTGCCGGTAGCAGCC**CCCTAACCCTAA**TGAGCGGTGGTTATCCA*G
M-3	5'CCTGGATAACCACCGCTCA**TTAGGGTTAGGG**ACTCAACTGCAGTCGCTT*G
M-4	5'CAAGCGACTGCAGTTGAGT**CCCTAACCCTAA**CCTTGCTAGGACGATGAT*T

3.3 RESULTS

3.3.1 WRN helicase and exonuclease activities simultaneously process Holliday Junctions

Prior to investigating TRF2 modulation of WRN activity on HJ substrates, we required a more complete understanding of WRN mechanism for HJ dissociation. Previous independent studies demonstrated WRN exonuclease activity at the 3' blunt end of a static 4-way junction (141) and WRN helicase dissociation of a 4-way junction with a mobile homologous core (30). Therefore, we asked whether both activities can simultaneously process a model HJ with a mobile core. To test this, we developed a multiplex fluorophore imaging method to track the fate of individual strands of a 4-way construct in with three strands were labeled with a unique fluorophore. Previous studies with 4-way junctions monitored substrate conversion to ssDNA by tracking migration of one radio-labeled strand in the construct. While current fluorophores still suffer the disadvantage of decreased sensitivity compared to ^{32}P radioactivity, they offer the distinct advantage of identifying the exact oligonucleotide composition of all intermediates and products present in each band on a gel. This is particularly critical if the HJ construct is asymmetric due to the presence of telomeric repeats or chemically modified bases in some arms, as in many of the HJs in this study (**Table 3.1**).

Each arm of the HJA construct (**Table 3.1**) is 25 bp long with a 12-nucleotide homologous center (30). Since a 4-way junction is unlikely to have four exposed ends *in vivo*, all the 3' ends except for the T-oligonucleotide contained a thiophosphate bond to protect against exonuclease activity. One exposed end mimics a regressed replication fork which is thought to be a target for WRN (134;142;146). WRN processing of the HJA substrate led to a variety of displaced DNA products that were uniquely identified by their fluorescent tags (potential products shown separately in Fig. 3.3). The appearance of a forked duplex preceded the

appearance of ssDNA product (**Fig. 3.2A**), as shown schematically in Fig. 3.2B. The distribution of the T-oligo, which contains the unprotected 3'end, among the various intermediates and products was quantitated (**Fig. 3.2C**). In the first eight minutes of the reaction less than 5% of the HJ was unwound and the predominant displaced T-oligo species was an intact fork duplex (3% of the total DNA) (**Fig. 3.2A**, lanes 2-10 and 3.2C). After 16 minutes, both T-oligo exonuclease degraded fork (exo-fork) and ssDNA (exo-ssDNA) species began to predominate (8% and 3%, respectively), and increased linearly with time to 19% and 18%, respectively, at 60 minutes (Fig. **3.2C**). After 60 minutes the amount of intact T-oligo in a fork or as ssDNA was low (< 3%) or negligible, respectively, indicating robust WRN exonuclease attack of the vulnerable arm during the reaction (Fig. **3.2D**). Upon completion the A-oligo, which pairs with both the T- and C-oligos in the HJ, was evenly divided between an intact fork with the C-oligo (8%) and a degraded fork with the T-oligo (11%; Fig. **3.2D**). Similar amounts of fork and ssDNA product were detected for each oligo, indicating no apparent bias for dissociation of any of the arms (Fig. **3.2D**). Our data are consistent with a mechanism for HJ processing in which WRN simultaneously displaces the HJ into forks while degrading the vulnerable T-oligo, followed by fork displacement to ssDNA.

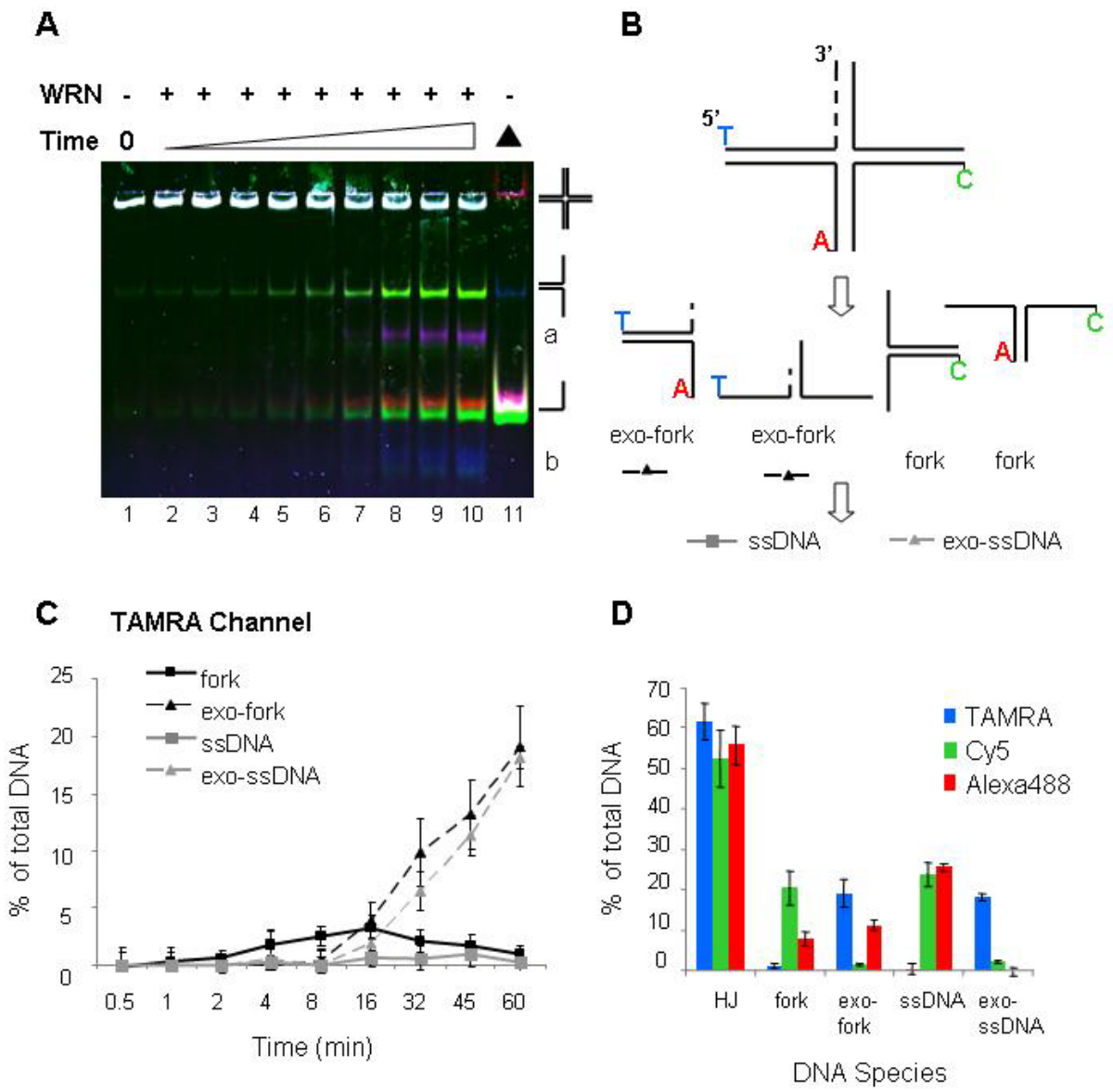

Fig. 3.2. WRN helicase and exonuclease simultaneously process HJ DNA. (**A**) HJA (2 nM) was incubated with 24 nM WRN in standard reaction buffer at 37°C. Aliquots were terminated at various times from 0.5 to 60 min in two-fold increments (lanes 2-10). Reactions were run on a 12% native polyacrylamide gel and were visualized with a Typhoon Imager in the TAMRA (blue), Cy5 (green), and Alexa 488 (red) channels. Colors assigned by Imagequant software. The exonuclease degraded products are indicated as (a) = exo-fork and (b) = exo-ssDNA. (**B**) Schematic of WRN strand displacement and exonuclease products. T = TAMRA, C = Cy5, and

A = Alexa488 labeled oligonucleotides. The 3' end of the exonuclease vulnerable T-oligo is highlight as a dotted line. (**C**) The percent of displaced T-oligo products from the reactions in (**A**) were quantitated as describe in **Materials and Methods** and plotted against time. Intact fork, solid black line and squares; exonuclease-degraded fork, dashed black line and triangles; intact ssDNA, solid gray line and squares; exonuclease-degraded ssDNA, dashed gray line and triangles. (**D**) Substrate and product distribution for each labeled oligonucleotide in the HJ construct after 1 h reaction. The percent of T-, A- and C-oligos present in the HJ substrate and each intermediate and product were quantitated as a function of total oligonucleotide as described in **Materials and Methods**. Values represent the means and standard deviations of two or three separate experiments.

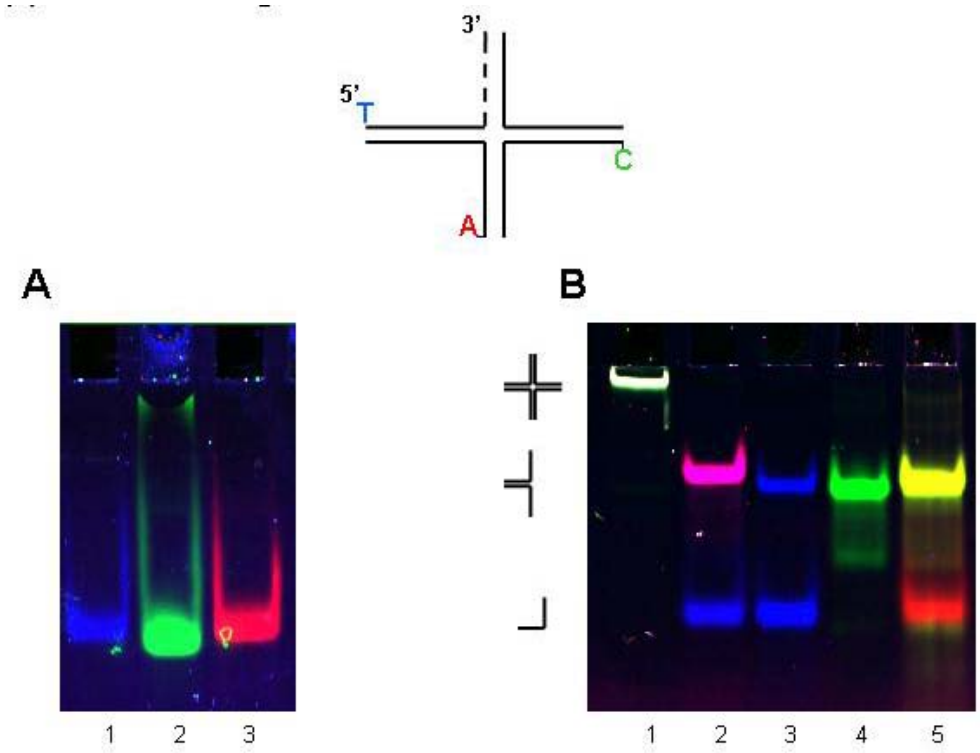

Fig. 3.3 Potential products from WRN helicase displacement of the HJA construct. Schematic of the HJA construct. T = TAMRA, C = Cy5, and A = Alexa488 labeled oligonucleotides. The 3' end of the exonuclease vulnerable T-oligo is highlight as a dotted line. Products were run on a 12% native polyacrylamide gel and were visualized with a Typhoon Imager in the TAMRA (blue), Cy5 (green), and Alexa 488 (red) channels. Colors assigned by Imagequant software. (**A**) T-oligo (lane 1), C-oligo (lane 2), A-oligo (lane 3). (**B**) HJA substrate (lane 1), T- and A- oligos in a forked duplex (lane 2), T- and unlabeled oligos in a forked duplex (lane 3), C- and unlabeled oligos in a forked duplex (lane 4) and C- and A- oligos in a forked duplex (lane 5).

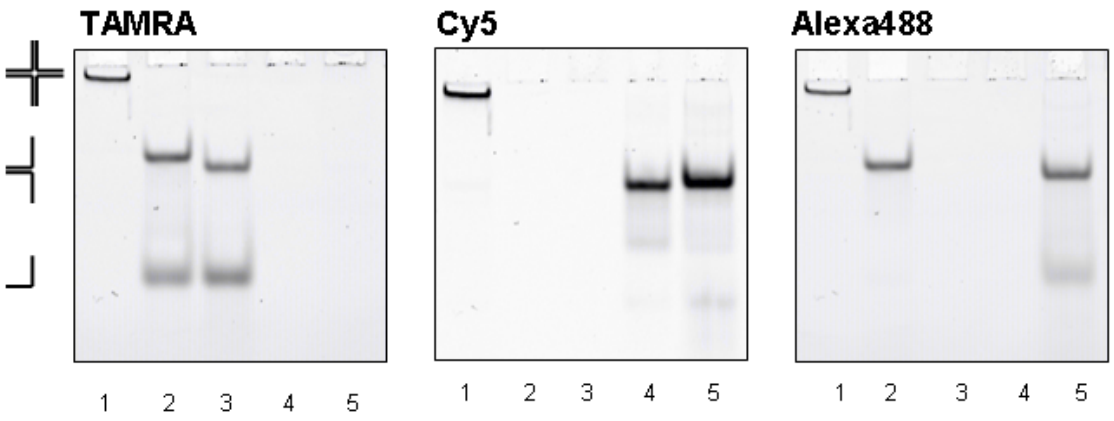

Fig. 3.4 Fluorophore emission demonstrates negligible crosstalk. **A**, TAMRA emission channel; **B**, Cy5 emission channel; **C**, Alexa488 emisison channel. HJA substrate (lane 1), T- and A-oligos in a forked duplex (lane 2), T- and unlabeled oligos in a forked duplex (lane 3), C- and unlabeled oligos in a forked duplex (lane 4) and C- and A- oligos in a forked duplex (lane 5).

3.3.2 WRN helicase initiates HJ strand displacement at the center

WRN binds the HJ core and is presumed to initiate unwinding from the center (142). However, while WRN helicase is inactive on blunt ended dsDNA, the WRN exonuclease can digest the blunt ends of substrates that contain junctions such as forks, bubbles or HJs, presumably by WRN first loading at the junction site (142). Therefore, it is formally possible that WRN helicase may also initiate strand displacement from the blunt end of an HJ, such as at a regressed replication fork, which may bypass effects of TRF2 bound at the core. To test the mechanism of HJ unwinding we used strategically placed biotin/streptavidin blocks. We and others showed that while a biotin does not interfere with WRN helicase activity (**Fig. 3.1**), a biotin/streptavidin complex can inhibit strand displacement (34). Control experiments confirmed that a biotin/streptavidin complex inhibits X-WRN unwinding of a forked duplex only when present on the translocating strand (**Fig. 3.5A**), but not when present on the non-translocating strand (**Fig. 3.5B**). Slower migration of the fork and ssDNA product upon streptavidin addition confirms the presence of a biotin/streptavidin complex. An exonuclease-dead variant of WRN (X-WRN) was used to examine helicase alone. Since WRN helicase is poorly processive, RPA was required to increase product yield for detection of potential inhibition, but did not relieve the biotin/streptavidin block to WRN unwinding and does not melt the duplex (**Fig. 3.10**).

We reasoned that if WRN initiates unwinding from the HJ center then a biotin/streptavidin block at the core, 24 nucleotides from the 5'end of the T-oligo (HJbio-center), would inhibit WRN 3' to 5' translocation on the T-oligo through the T/A duplex arm (**Fig. 3.6A** III, open arrows). However, unwinding in the "vertical" axis (dark arrows) would be permitted (**Fig. 3.6A**, dark arrows) yielding products that included the T-oligo trapped in a fork with the A-oligo. When HJbio-center was incubated with X-WRN and RPA, 47% and 89% of the T- and A-

oligos were liberated as ssDNA, respectively (Figs **3.6B, C** lane 2, and **3.6D**). Some of the T-oligo was released as a fork with the unlabelled oligo, but not with the A-oligo, perhaps due to greater G/C content on the HJ arm composed of the T-oligo and the unlabeled strand (Table 3.2). Upon streptavidin addition, the displaced ssDNA product decreased substantially to 17% and 55% for the T- and A-oligos, respectively (Figs **3.6B, C** lane 5, and **3.6D**). Importantly, streptavidin addition yielded a novel forked species consisting of the T- and A-oligos (total of 31% and 15%, respectively, since a T/unlabeled fork is also present) and a novel triple stranded species consisting of the T- and A-oligos (17% and 15%, respectively) bound to the unlabeled oligo. These species contained negligible amounts of C-oligo (1%; **Fig. 3. 7**) and the 3-way probably resulted from shifting between two conformations (**Fig. 3.6E**). The 12 nucleotide homologous core can shift the biotin/streptavidin block from the T/A duplex arm (**Fig. 3.6A**, I and II) to the T/unlabeled duplex arm (**Fig. 3.6A**, III), whereas the C-oligo is never present in a duplex arm with a biotin (**Fig. 3.7A**). These dynamic conformations likely explain the lack of complete WRN inhibition and the generation of a novel 3-way species by a biotin/strepativin barrier at the HJ core.

To test whether WRN can load on the HJ blunt end and translocate toward the center while unwinding, we asked if a 3' biotin/streptavidin complex on the A-oligo would inhibit WRN helicase activity (HJbio-end) (**Fig. 3.6A**). In contrast to HJbio-center, streptavidin addition to HJbio-end did not alter strand displacement by X-WRN and RPA. Approximately 95% of the A-oligo was displaced as ssDNA in the absence or presence of streptavidin (**Fig. 3.6A**, lanes 2-4 or 7, respectively), and the 3-way product observed in the HJbio-center reactions was not apparent. Furthermore, RPA alone does not melt the HJ (**Fig. 3.16**). These data indicate that a

biotin/streptavidin complex at the HJ core, but not at the HJ blunt end, inhibits WRN strand displacement.

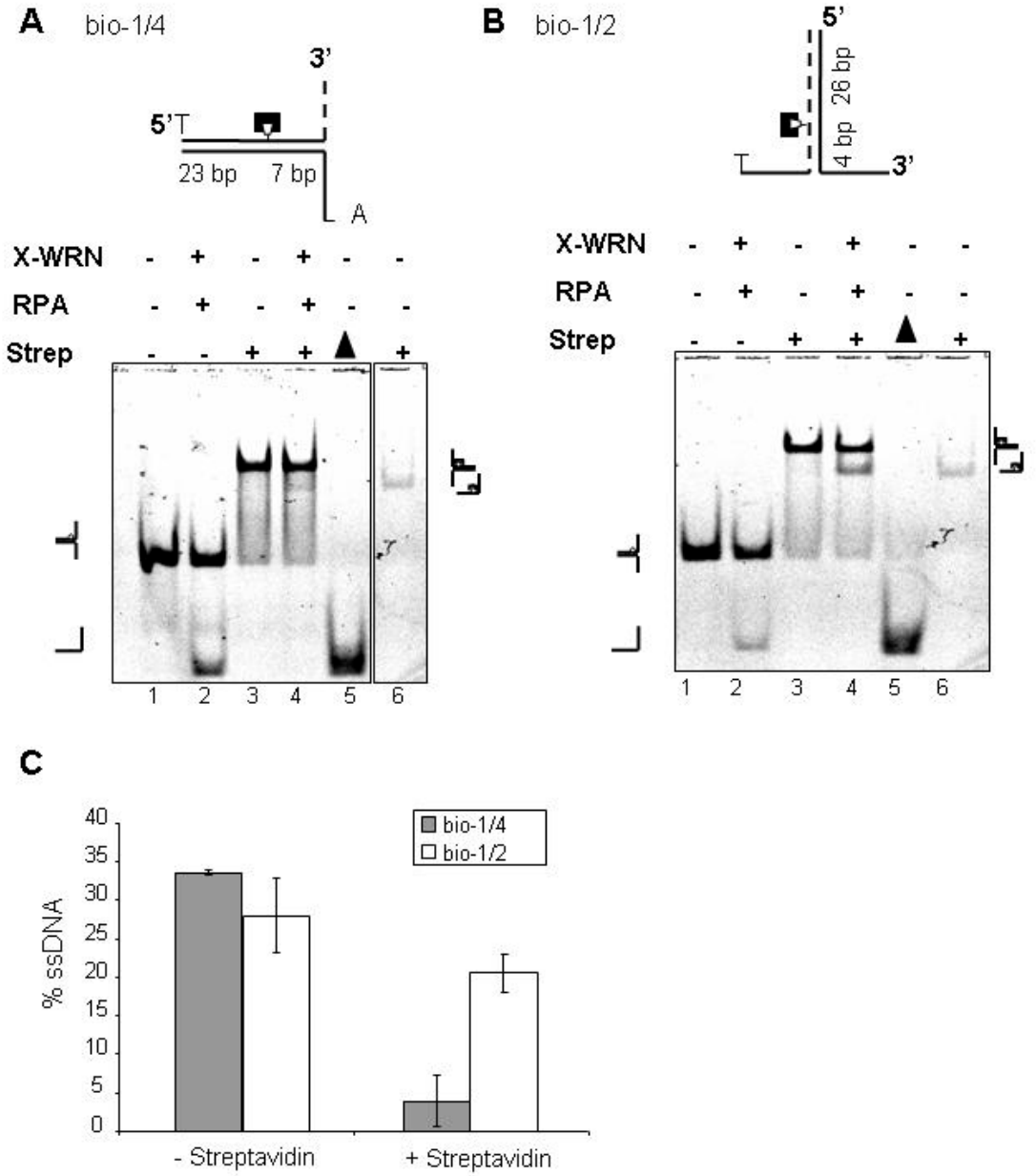

Fig. 3.5 Biotin-streptavidin complex on the translocating strand inhibits WRN activity. Reactions contained a 31-bp forked duplex (0.5 nM) with a biotinylated nucleotide (inverted trapezoid) on the WRN translocating (**A**) or nontranslocating strand (**B**). Black box denotes streptavidin. The forks were constructed by annealing oligonucleotides bio-1 and 4 (**A**) or bio-1 and 2 (**B**) (Supplemental Table 1). T = TAMRA; A= Alexa488. The duplex length (bp) on either side of the biotinylated nucleotide is shown. The substrate was pre-incubated with either 0 (lanes 1-2) or 30 nM (lanes 3-4) streptavidin in standard reaction buffer prior to adding 3.4 nM X-WRN and

12 nM RPA. Reactions were for 20 min at 37^{o}C, and were run on an 8% native polyacrylamide gel for 2.5 hrs and visualized with a Typhoon imager. A streptavidin-bound T-oligo was loaded (lane 6) as a marker and the boiled substrate lane is indicated with a triangle (lane 5). The gel scans show the TAMRA emission channel**.** (**C**) Quantitation of the ssDNA product. The percent of ssDNA T-oligo product was quantitated as described in **Materials and Methods**. Values represent the means and s.e.m. from two independent experiments.

Figure 3.6

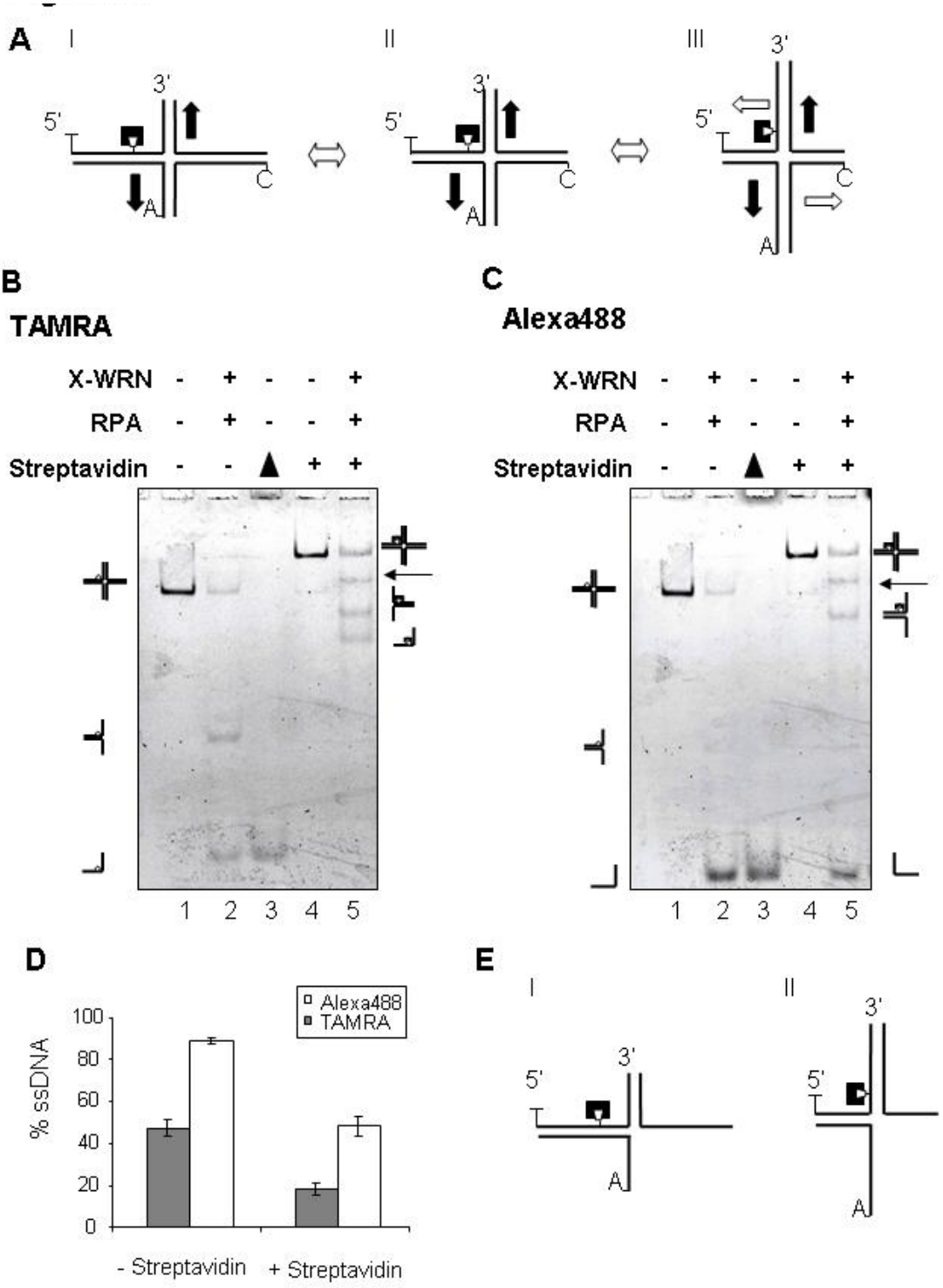
A
I
II
III
3'
5'
C
A
B
TAMRA
C
Alexa488
X-WRN - + - - +
RPA - + - - +
Streptavidin - - ▲ + +
1 2 3 4 5
D
Alexa488
TAMRA
% ssDNA
100
80
60
40
20
0
- Streptavidin
+ Streptavidin
E
I
II

Fig. 3.6 A biotin-streptavidin complex at the HJ core impedes WRN helicase activity. (**A**) The HJbio-center construct has a 12-nucleotide homologous core that permits branch migration. The center construct (II) shows the HJ in a symmetric conformation with 25 bp arms. The left and right constructs (I and III) show the possible extreme conformations. T = TAMRA; C = Cy5; A = Alexa488; inverted trapezoid denotes a biotin moiety. Black box denotes a streptavidin. The predicated possible directions of unwinding are shown for each conformation; white arrow denotes unwinding along the horizontal axis; black arrow denotes unwinding along the vertical axis. (**B** and **C**) The HJbio-center substrate (0.5 nM) was pre-incubated without (lanes 1-2) or with (lanes 4-5) 30 nM streptavidin under standard reaction conditions prior the addition of 15 nM X-WRN and 45 nM RPA. The reactions were conducted for 1 h at 37^{o}C. The products were run on an 8% native acrylamide gel for 2.5 hrs and visualized with a Typhoon Imager. Black triangle denotes boiled substrate lane. The TAMRA and Alexa488 emission channels are shown in (**B**) and (**C**), respectively. (**D**) Quantitation of ssDNA reaction products. The percent of T-oligo (grey bars) and A-oligo (white bars) detected as ssDNA product was quantitated as described in **Materials and Methods**. The values represent the means and standard deviations from two to four independent experiments. (**E**) Schematic of the 3-way product species that are generated upon addition of streptavidin.

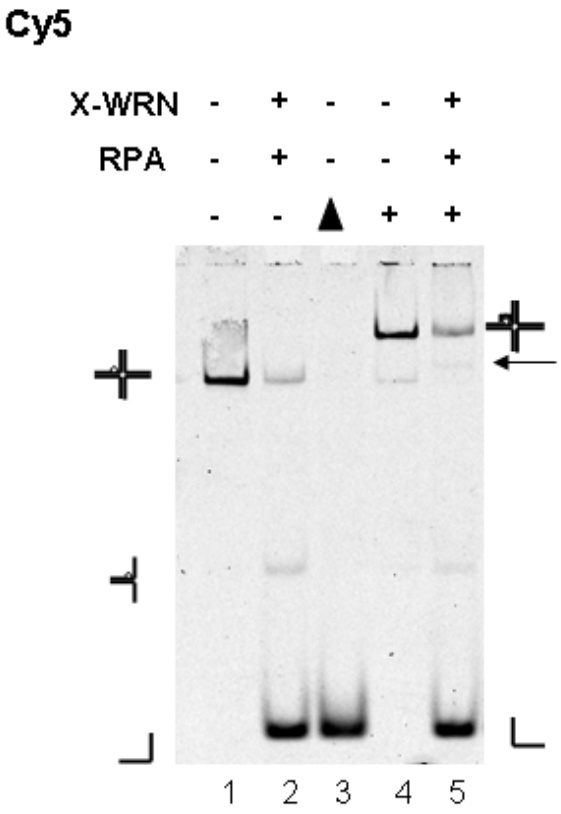

Fig. 3.7 Streptavidin-biotin barrier fails to create a triple-stranded species with C-oligo

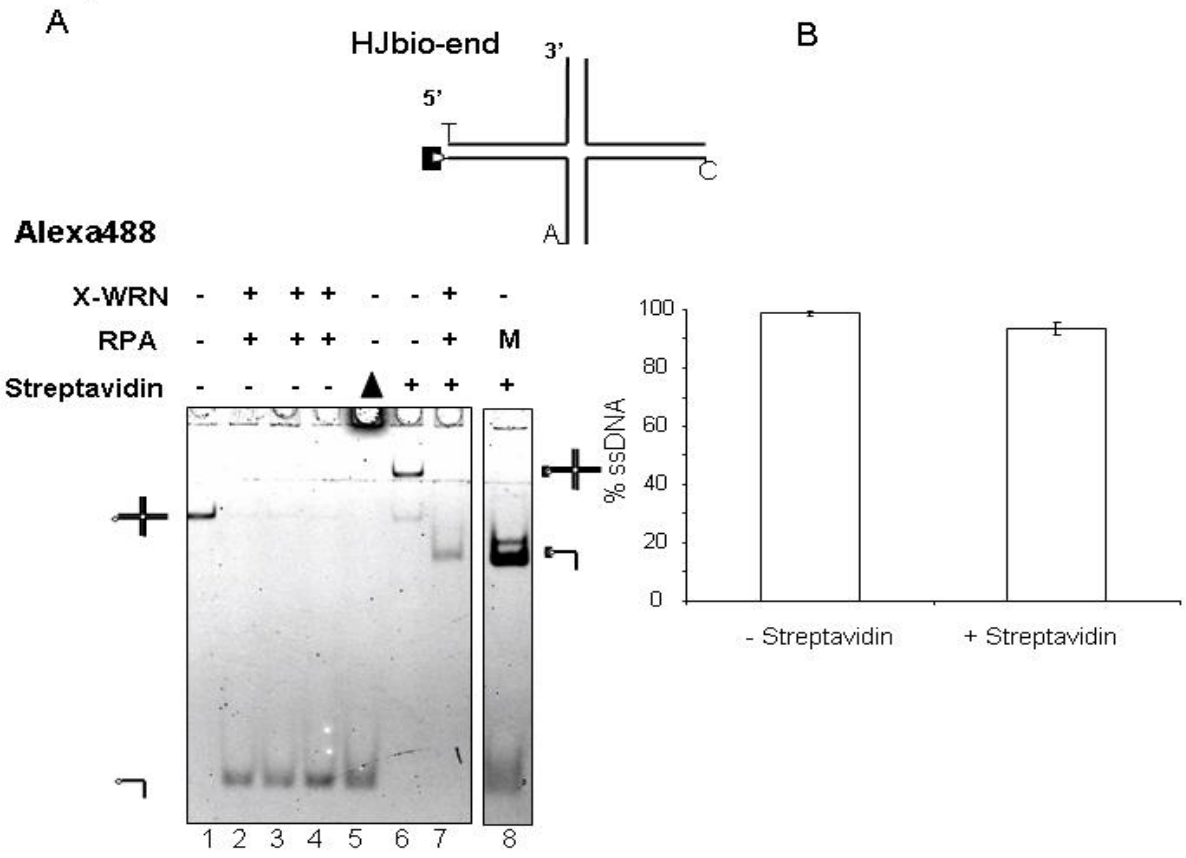

Fig. 3.8 A biotin-streptavidin complex at the 3' end of an HJ arm does not alter WRN activity. The HJbio-end substrate has a biotin tag at the 3' end of the A-oligo. T = TAMRA; C = Cy5; A = Alexa488; inverted trapezoid denotes a biotin moiety. Black box denotes streptavidin. (**A**) The HJbio-center (0.5 nM) substrate was pre-incubated without (lanes 2-4) or with (lanes 6-7) 30 nM streptavidin under standard reaction conditions prior the addition of 15 nM X-WRN and 45 nM RPA. The reactions were conducted for 1 h at 37^{o}C. The products were run on an 8% native acrylamide gel for 2.5 hrs and visualized with a Typhoon Imager. Black triangle denotes boiled lane. M denotes an A-oligo ssDNA marker bound by streptavidin. The Alexa488 emission channel is shown. (**B**) Quantitation of ssDNA reaction products. The percent A-oligo detected as

ssDNA product was quantitated as described in **Materials and Methods**. The values represent the means and standard deviations from three independent experiments.

Prior to the biotin-streptavidin barriers, we had experimented with PEG linkers to interrupt the DNA backbone, which would also create a barrier to WRN helicase. We found that if the PEG linker interrupted the translocating strand of a forked duplex (Peg-1), the X-WRN unwinding was severely inhibited, unwinding only 1% of the forked substrate at a 0.75:1 ratio of RPA to X-WRN, and unwinding 9% at a 3:1 RPA:X-WRN ratio (**Fig. 3.9A, C**). By contrast, if the PEG linker interrupted the complement of the strand upon which X-WRN translocated (Peg-2), there was six-fold more unwinding at the 3:1 RPA:X-WRN ratio compared to when the PEG was in the translocating strand. (**Fig. 3.9B,C**).

We then tested the effect of PEG linkers on a Holliday Junction construct in order to determine where WRN helicase initiates unwinding of the HJ DNA. We constructed two HJs identical to HJA (Table 3.1) but for the replacement of three nucleotides on the TAMRA-labeled oligonucleotide with a PEG linker on the 5' or 3' half of the oligonucleotide (HJ-5' and HJ-3', respectively; **Table 2.1** ; **Fig. 3.10A,B**, respectively). If WRN helicase unwinds HJ from the center and progresses outward, then we predict WRN would not unwind HJ-5', but would unwind HJ-3'. We observed that HJ-5' was poorly unwound by X-WRN alone under standard conditions, with more than 90% of the HJ remaining (**Fig. 3.10C**) after one hour. This was half the amount of unwinding seen with X-WRN alone with the HJ-3' construct (**Fig. 3.10D**), where the PEG linker is on the 3' half. Because at most only 20% of the PEG-containing HJ constructs were unwound by X-WRN alone after one hour, we used RPA to produce more unwinding to enhance the ability to detect a potential difference in unwinding between the two constructs. With the addition of RPA in a 3:1 molar ratio over X-WRN, HJ-5' unwinding incrased nearly 10-fold, with 50% more HJ-3' A-oligo being unwound than HJ-5'(**Fig. 3.10C**). The relative inhibition of X-WRN-mediated unwinding of the HJ with the PEG at the 5'end of the TAMRA

strand, compared to the HJ with the PEG at the 3’ end of the TAMRA strand, is consistent with a model in which WRN initiates unwinding from the center of the HJ.

To confirm that WRN helicase does not initiate unwinding from the end of HJ DNA, we created a 5' PEG HJ DNA (HJ-5bio) construct as above with a 3' biotin tag on the 3' end of the Alexa488-labeled fluorophore (**Fig. 3.11**; **Table 3.1**). Biotin-streptavidin complexes have been used to block WRN helicase processivity and thus determine mechanisms of unwinding different substrates(147). We hypothesized that if WRN helicase initiated unwinding from the center of the Holliday Junction, then WRN unwinding should not be significantly changed by the presence or absence of streptavidin. However, if WRN helicase initiated unwinding from the edge of the HJ and moved inward, the presence of streptavidin should block WRN activity. When we reacted HJ-5bio with 3:1 RPA:X-WRN under standard conditions, we found that the presence or absence of streptavidin in a 60:1 ratio with DNA did not significantly alter unwinding (**Fig. 3.11A**), with 44% HJ DNA unwound without strepavidin, and 48% being unwound with streptavidin (**Fig. 3.11C**). The replacement of basepaired DNA with the linker probably destabilized the HJ, especially since branch migration could result in the PEG linker being only 9 basepairs from the center. A biotin-streptavidin barrier does not remove basepairs, thus creating a more stable construct.

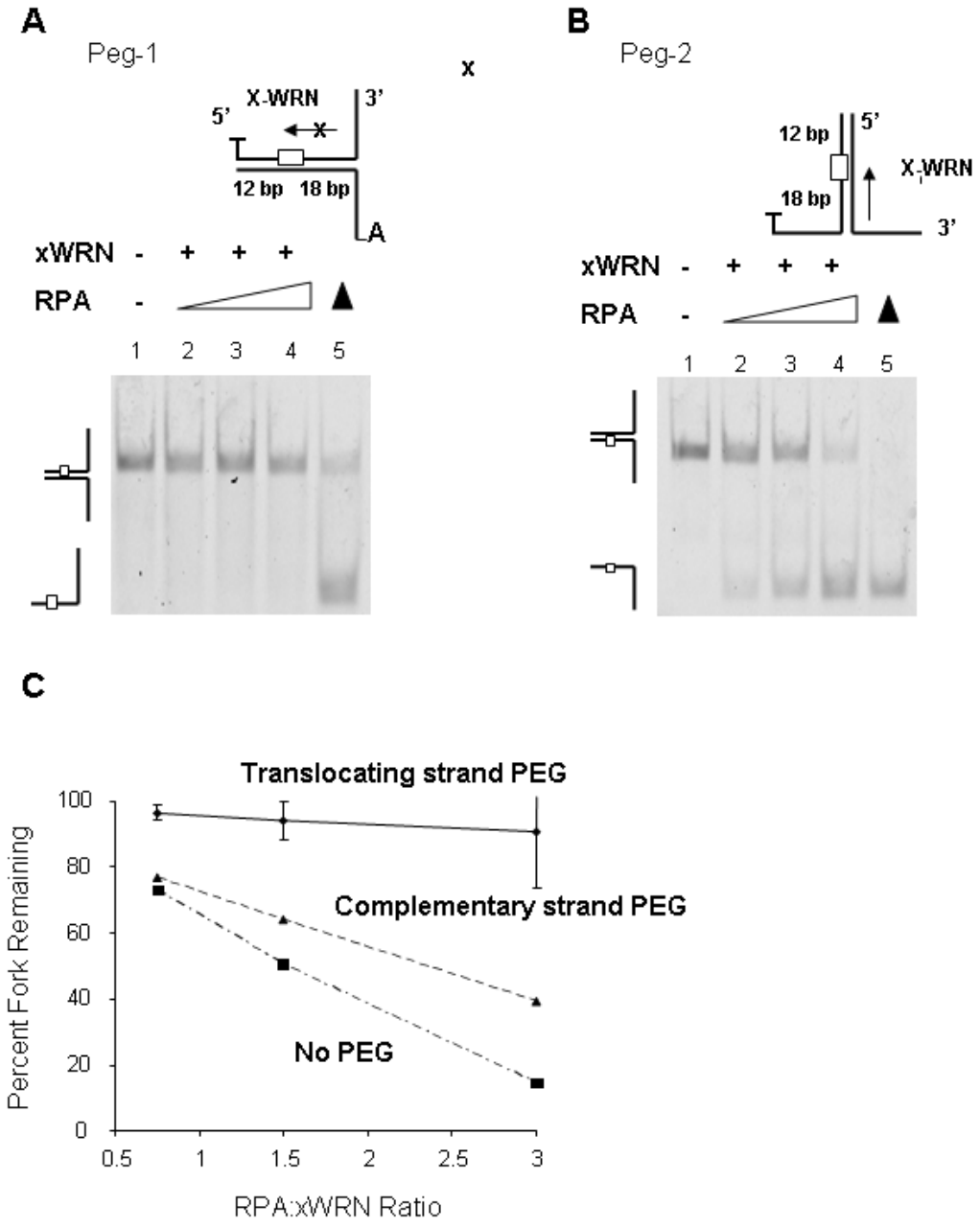

Fig. 3.9 A PEG linker on the translocating strand inhibits an X-WRN-RPA complex. Reactions contained 1 nM of a forked 31 bp duplex DNA with 3 nucleotides of either the translocating (**A**), or non-translocating (**B**) strand replaced by a PEG linker, represented by an open box, causing a discontinuity in the phosphate backbone. The duplex was incubated with 11.25, 22.5, or 45 nM RPA and 15 nM X-WRN (lanes 2-4) for 20 min. under standard reaction conditions as described

in **Materials and Methods**. The reactions were stopped with 11.5 µL ProtK-LiCl stop solution and precipitated as described in **Materials an d Met hods** for the PEG-containing Holliday Junctions. The circle denotes a TAMRA dye, and the diamond an Alexa 488 dye. **C.** The percent remaining forked DNA for **A** and **B**, along with a control forked substrate that has no PEG linker and reacted under identical conditions to **A** and (**B**), graphed versus the ratio of RPA to X-WRN.

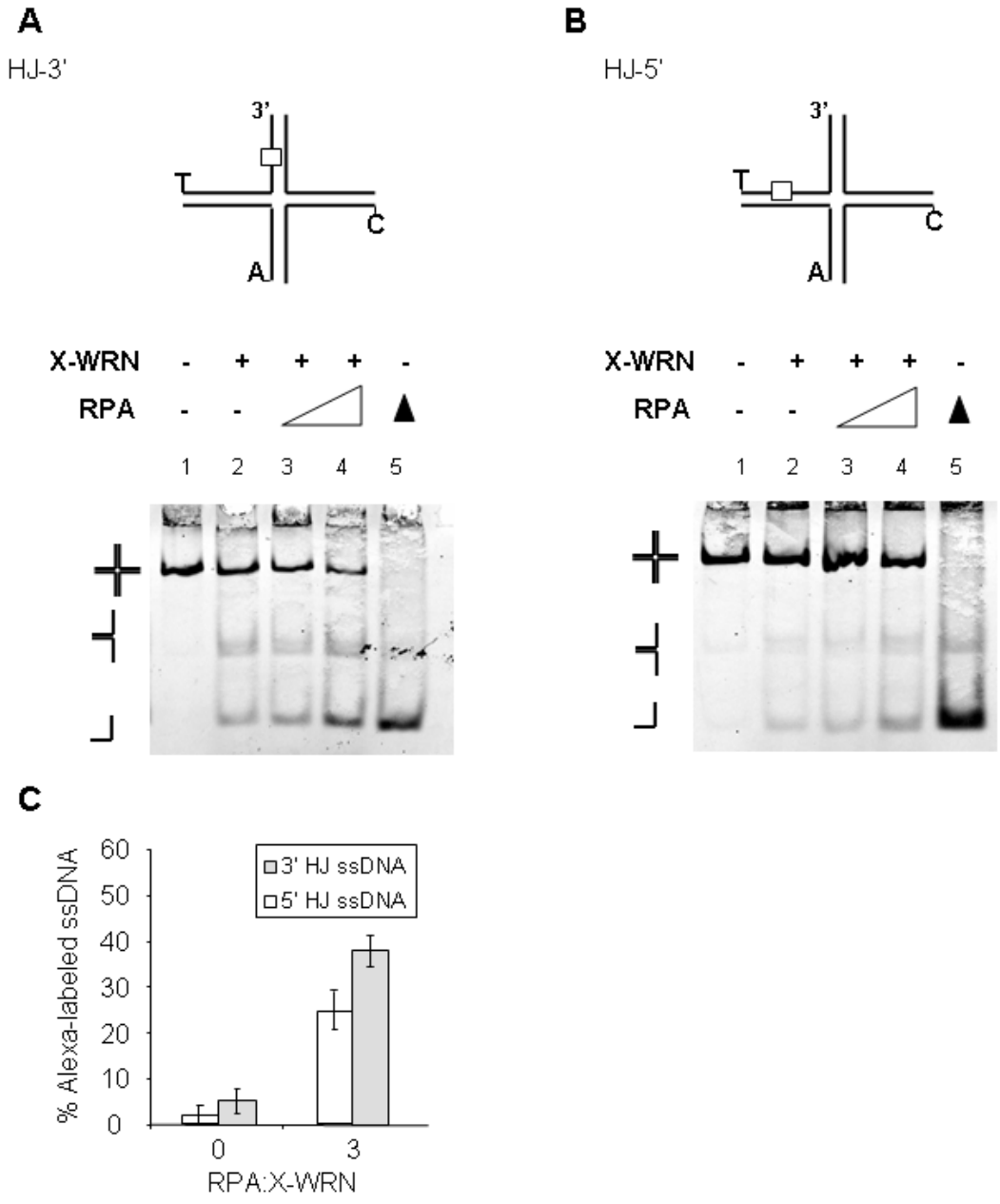

Fig. 3.10. A 5' PEG barrier on a Holliday Junction partially inhibits strand displacement. T denotes TAMRA fluorophore; C denotes Cy5 dye, A denotes Alexa488 dye. Black triangle denotes boiled lane and an open rectangle denotes a PEG linker. 2 nM HJbio-3' (**A**) or 2 nM HJbio-5' (**B**) is reacted with 15 nM X-WRN and 0, 22.5, or 45 nM RPA (lanes 2-4) for 60

minutes. The reactions are stopped by 11.5 µL ProtK-LiCl stop solution and precipitated as described in **Materials and Methods** The gel was then imaged as described above, and the Fig. shows the Alexa488 emission channel. **C**, quantitation of the HJbio-3' and HJbio-5' reactions in (**A**) and (**B**) as a function of RPA. Displaced helicase products are calculated as a mean percentage of the total DNA in the lanes versus the RPA:X-WRN ratio (see **Materials and Methods**). The means and standard deviations are calculated from two to four independent reactions.

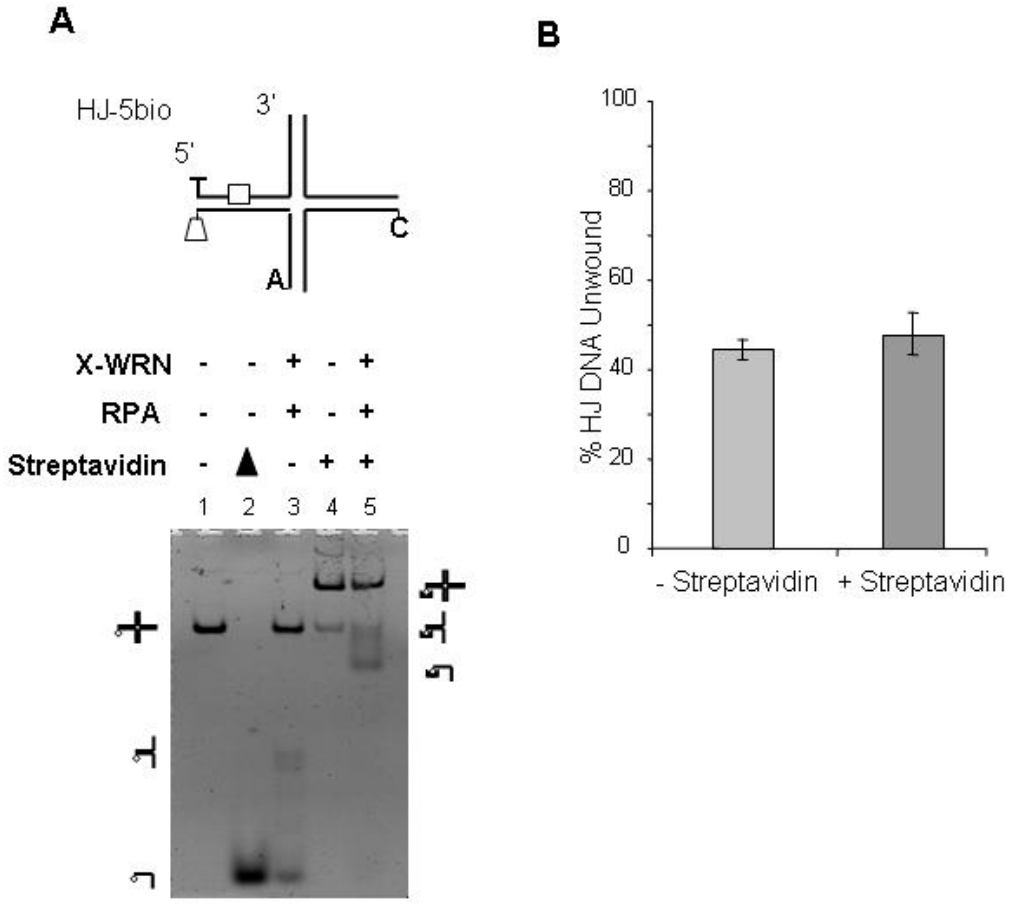

Fig. 3.11 Biotin-streptavidin does not further impede WRN on the 3' end of HJ-5'. T denotes TAMRA fluorophore; C denotes Cy5 dye, A denotes Alexa488 dye and a trapezoid denotes a biotin tag. Black triangle denotes boiled lane And open rectangle denotes a PEG linker. (**A**) 2 nM HJ-5bio is reacted with 15 nM X-WRN and 45 nM RPA, with (lane5) or without (lane 3) 120 nM streptavidin, under standard reaction conditions, and stopped by 11.5 µL ProtK-LiCl stop solution and precipitated as described in **Materials and Methods**. The gel was then imaged as described above, and the Fig. shows the Alexa488 emission channel. (**B**) Quantitation of the HJ-

5bio reactions in **A**. Displaced helicase products are calculated as a mean percentage of the total DNA in the lanes versus the RPA:X-WRN ratio (see **Materials and Methods**). The means and standard deviations are calculated from two to four independent reactions.

3.3.3 TRF2 protects telomeric arm HJ DNA from WRN strand displacement activity

We hypothesized that TRF2 protects HJs from WRN activity since it stabilizes HR-like structures in the telomere from unwanted processing *in vivo* (1). To test this we examined the non-telomeric HJA, and designed an HJ containing three telomeric repeats in two duplex arms in a *trans* configuration (HJT; **Table 3.1**). A fully telomeric HJ construct would have been unstable due to spontaneous branch migration. TRF2 pre-binding to the substrate decreased the percent of WRN HJT strand displacement in a dose-dependent manner from 49% to 20% (**Fig. 3.12A,D**). A TRF2 mediated decrease in exonuclease-degraded products was also observed (**Fig. 3.13**). To determine whether TRF2 binding to the HJ core via its B-domain was responsible for WRN inhibition, we tested the non-telomeric HJA which lacks Myb binding sequences. In sharp contrast to the telomeric arm HJT, pre-binding TRF2 to HJA did not alter WRN unwinding (Fig. **3.12B and D**).

Next, we asked whether TRF2 interaction with the HJ arms via the Myb domain was responsible for TRF2 inhibition of WRN activity on the telomeric HJT. To test this, a TRF2 fragment with the N-terminal B domain deleted (TRF2ΔB) was used. The pre-incubation of telomeric HJT with TRF2ΔB exhibited no protective effect against WRN strand displacement activity (**Fig. 3.12C** and **D**). These experiments indicate that TRF2 protects telomeric HJ DNA from WRN activity, but only on substrates containing binding sites for both the TRF2 B and Myb domains.

A previous study reported that TRF2 protects an HJ with two telomeric repeats in the center from cleavage by HJ resolvase enzymes (84). We tested the identical construct (HJM, **Table 1**) to determine whether TRF2 could provide a comparable protection from WRN strand displacement. While HJM has binding sites for both the TRF2 Myb and B domains, the Myb

domain can only bind when the homologous core migrates to an extreme conformation with the two telomeric repeats in tandem on the same duplex arm (**Fig. 3.13A**). In contrast to telomeric arm HJT, the pre-incubation of telomeric core HJM with increasing TRF2 concentrations did not alter the WRN strand displacement or exonuclease products (**Fig. 3.13B,D**). To determine whether the lack of TRF2 inhibition resulted from WRN digestion into the Myb binding sequences, we repeated the experiment with the exonuclease-dead X-WRN. However, TRF2 did not alter X-WRN strand displacement of HJM either (**Fig. 3.13C**, **D**).

3.3.4 RAP1 does not alter TRF2 protection of telomeric arm HJ from WRN activity

RAP1 is a member of the shelterin complex that binds TRF2, and is required along with TRF2 to protect telomeric ends from non-homologous end joining (NHEJ) (58;148). Therefore, we asked if RAP1 also cooperates with TRF2 to prevent telomeric HJ displacement. To test this, we pre-bound telomeric arm HJT with 25 nM TRF2 and increasing RAP1 concentrations (0 to 50 nM) prior to WRN addition (**Fig. 3.14B,** lanes 2-5). While TRF2 inhibition of WRN was apparent, no significant difference in the amount of displaced HJ DNA was observed as a function of RAP1 concentration (**Fig. 3.14C**). Nor did we detect any modulation of WRN activity by RAP1 (50 nM) alone (**Fig. 3.14B**, lane 6).

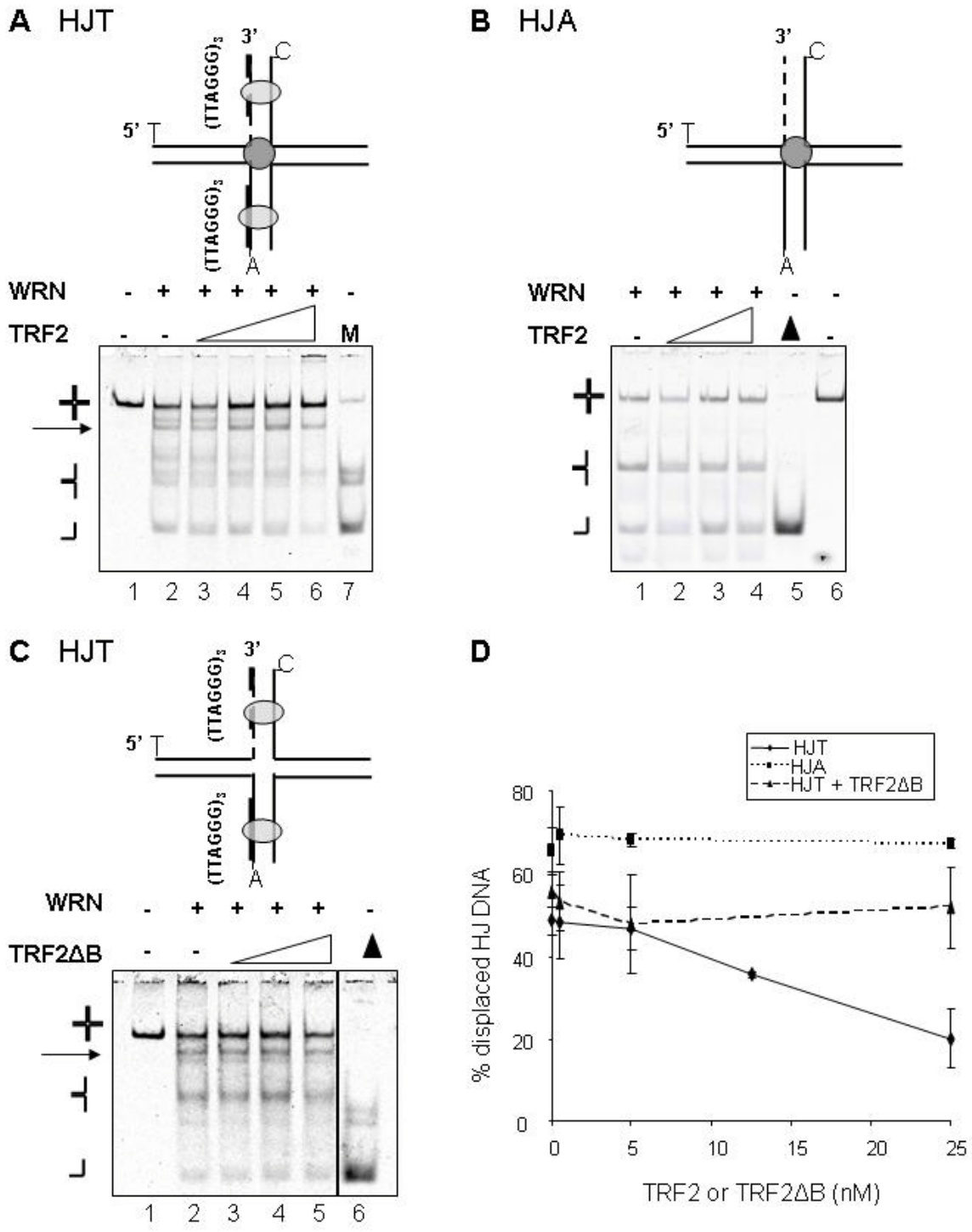
A HJT
3'
C
(TTAGGG)3
5'
(TTAGGG)3
A
WRN - + + + + + -
TRF2 - - M
1 2 3 4 5 6 7
B HJA
3'
C
5'
A
WRN + + + + - -
TRF2 - -
1 2 3 4 5 6
C HJT
3'
C
(TTAGGG)3
5'
(TTAGGG)3
A
WRN - + + + + -
TRF2ΔB - -
1 2 3 4 5 6
D
HJT
HJA
HJT + TRF2ΔB
% displaced HJ DNA
0 20 40 60 80
0 5 10 15 20 25
TRF2 or TRF2ΔB (nM)

Fig. 3.12 TRF2 inhibits WRN unwinding of HJ substrates with telomeric arms. A schematic of the telomeric arm (HJT) and non-telomeric HJA constructs are shown in panels (**A-C**). Thick black lines denote $(TTAGGG)_3$ repeats. T = TAMRA; C = Cy5; A = Alexa488. The 3' end of the exonuclease-vulnerable T-oligo is highlighted as a dotted [Legend continues on next page.] line. Predicted TRF2 binding sites are indicated with the circle (B-domain) and the ellipse (Myb domain). The substrates (0.5 nM) HJT (**A**) or HJA (**B**) were pre-incubated with either 0, 0.5, 5, 12.5 or 25 nM TRF2 (**A**, lanes 2-6, respectively) and either 0, 0.5, 5 or 25 nM TRF2 (**B**, lanes 1-4) in standard reaction buffer. The reactions were initiated by the addition of 19 nM WRN and reacted for 1 hr at 37°C. The reactions were run on 8% native gels for 1.5 hrs and visualized with a Typhoon Imager. Scans of the Cy5 emission channel are shown in panels (**A**-**B)**. Triangle indicates boiled substrate; M denotes a marker lane; arrows denote a triple-stranded species. (**C**) TRF2ΔB does not alter WRN unwinding of the telomeric HJ. The HJT (0.5 nM) substrate was pre-incubated with either 0, 0.5, 5 or 25 nM TRF2ΔB (lanes 2-5) prior to the addition of 19 nM WRN. Reactions were for 1 hr at 37°C, and were run on an 8% native gel. The scan from the Cy5 emission channel is shown. (**D**) Quantitation of HJ unwinding. The percent of C-oligo present in the HJ substrate was quantitated as a function of total C-oligo DNA in the reaction as described in **Materials and Methods** and plotted against TRF2 or TRF2ΔB concentration. The values represent the means and standard deviations from at least three independent experiments.

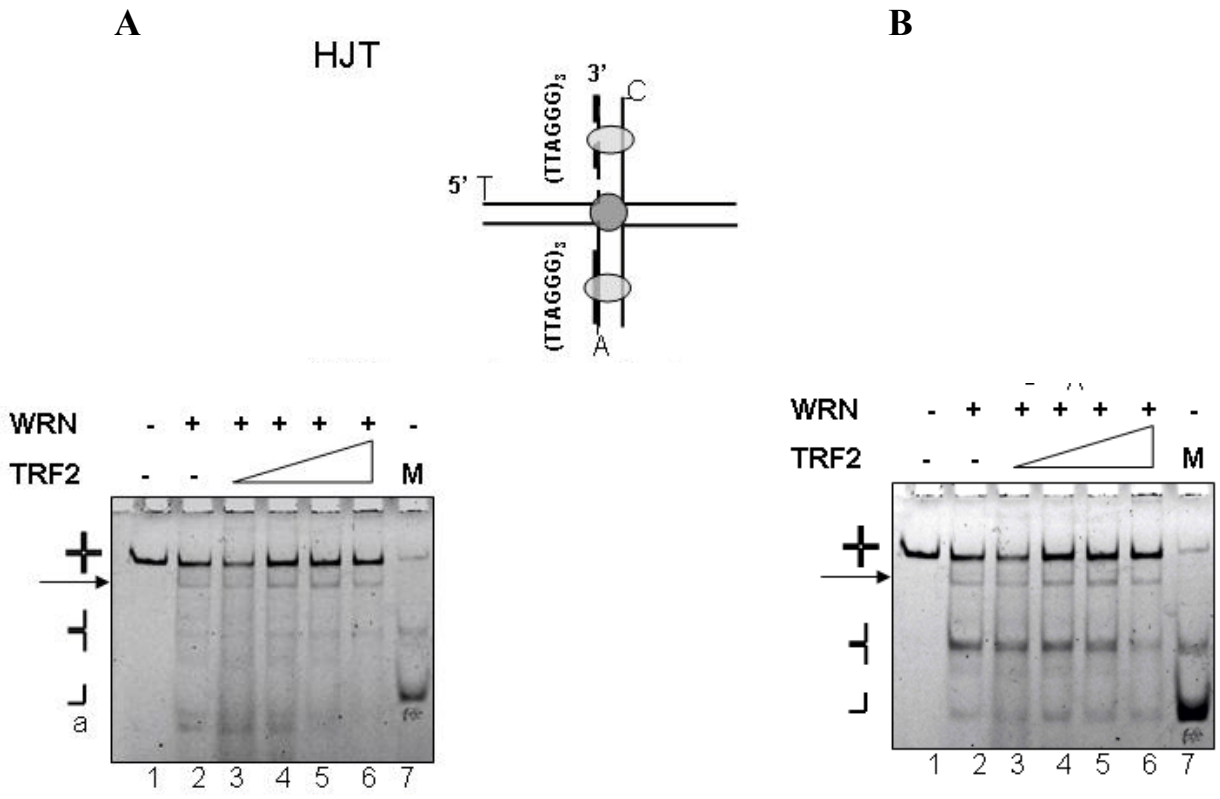

Fig. 3.13 TRF2 inhibits displacement of T- and A- oligos. Thick black lines denote $(TTAGGG)_3$ repeats. T = TAMRA; C = Cy5; A = Alexa488. The 3' end of the exonuclease vulnerable T-oligo is highlight as a dotted line. Predicted TRF2 binding sites are indicated with the circle (B-domain) and the ellipse (Myb domain). The substrate was pre-incubated with either 0, 0.5, 5, 12.5 or 25 nM TRF2 (lanes 2-6, respectively) in standard reaction buffer. The reactions were initiated by the addition of 19 nM WRN and reacted for 1 hr at 37^{o}C. The reactions were run on 8% native gels and visualized with a Typhoon Imager. **A**, TAMRA emission channel. **B**, Alexa488 emission channel. M denotes a marker lane; arrows denote a triple-stranded species; "a" denotes exonuclease-degraded ssDNA. the two telomeric repeats in tandem on the same duplex arm (**Fig. 3.8A**).

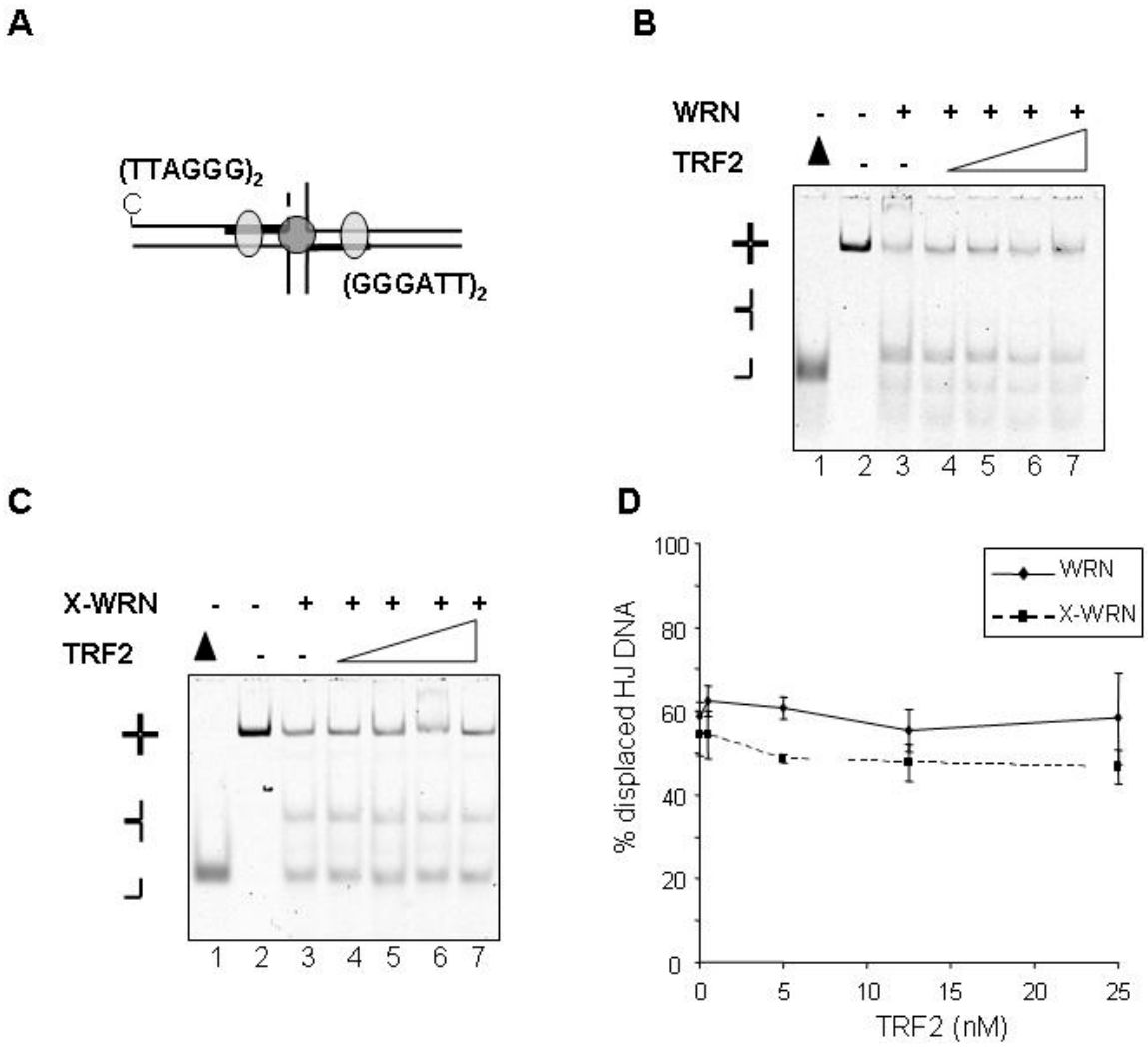

Fig. 3.14 TRF2 fails to inhibit WRN on an HJ with core telomeric repeats. (**A**) The HJM substrate has a 12 nucleotide homologous core composed of two (TTAGGG)2 repeats, denoted by a thick black line. Shown is the conformation with a fully branch migrated core. The predicted TRF2 binding sites are indicated with the circle (B-domain) and the ellipse (Myb domain). C = Cy5 fluorophore. The 3' end of the exonuclease vulnerable C-oligo is highlighted as a dotted line. (**B** and **C**) The HJM substrate (0.5 nM) was pre-incubated with either 0, 0.5, 5, 12.5 or 25 nM TRF2 (lanes 3-7) and reacted under standard conditions for 1 hour with 25 nM WRN (**B**) or 40 nM X-WRN (**C**), respectively. The reactions were run on 8% native gels for 1.5 hrs and visualized with a Typhoon Imager. The scan from the Cy5 emission channel is shown.

Triangle indicates boiled substrate**. (D**) Quantitation of HJ unwinding (% Cy5 oligonucleotide displaced).

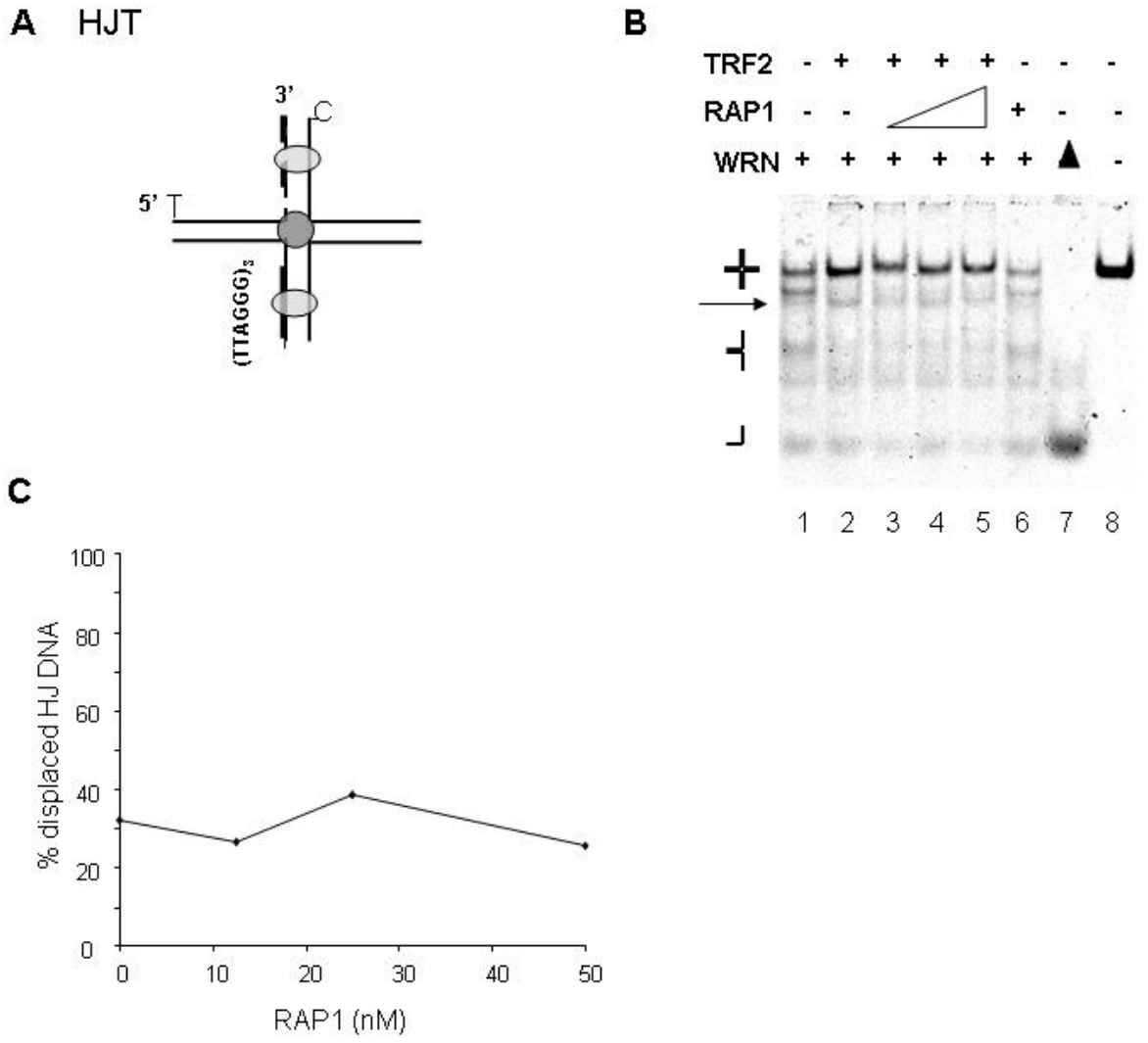

Fig. 3.15 RAP1/TRF2 complex inhibits WRN like TRF2 alone (**A**) Schematic of the HJT construct is shown. Thick black lines denote $(TTAGGG)_3$ repeats. T = TAMRA; C = Cy5; A = Alexa488. The 3' end of the exonuclease vulnerable T-oligo is highlight as a dotted line. Predicted TRF2 binding sites are indicated with the circle (B-domain) and the ellipse (Myb domain). (**B)** The substrate was pre-incubated with 25 nM TRF2 and either 0, 12.5, 25, or 50 nM RAP1 (lanes 2-5) in standard reaction buffer. The reactions were initiated by the addition of 19

nM WRN (lanes 1-6) and reacted for 1 hr at 37^{o}C. Lane 6 contained 50 nM RAP1 and 19 nM WRN. The reactions were run on 8% native gels and visualized with a Typhoon Imager. Scans of the Cy5 emission channel are shown. Triangle indicates boiled substrate; arrows denote a triple-stranded species. (**C**) Quantitation of HJ unwinding. The percent of C-oligo present in the HJ products was quantitated as a function of total C-oligo DNA as described in **Materials and Methods** and plotted against RAP1 concentration. Products were quantitated as a function of total C-oligo DNA as described in **Materials and Methods** and plotted against the TRF2 concentration. The values represent the means and standard deviations from three separate experiments.

3.3.5 TRF2 exhibits increased binding to telomeric HJ compared to non-telomeric HJ

We hypothesized that the mechanism by which TRF2 protects HJ DNA from WRN activity is via binding both the HJ core and the telomeric repeats in the HJ arms. Therefore, we predicted that TRF2 should exhibit increased binding to a telomeric HJ, compared to a non-telomeric HJ, since both TRF2 B and Myb domains can engage the telomeric HJ. We performed electrophoretic mobility shift assays to compare TRF2 binding to non-telomeric HJA and telomeric HJT and HJM. The migration of all three substrates was retarded by TRF2 in a protein concentration dependent manner (**Fig. 3.17A**, lanes 1-15). TRF2 shifted HJT to a higher migrating species compared to HJA or HJM, suggesting that more TRF2 molecules may be bound to the telomeric arm HJT substrate (**Fig. 3.17A**, lanes 5, 10, and 15). A shifted HJ/TRF2 complex was apparent at 50 nM TRF2 for HJT and HJM, but was not detected until 125 nM TRF2 for HJA (**Fig. 3.17A**, lanes 3, 8, and 13). The binding curves for the HJT and -M substrates showed a 50% shift by about 50 nM TRF2, compared to about 100 nM for HJA (Fig. **3.17C**).

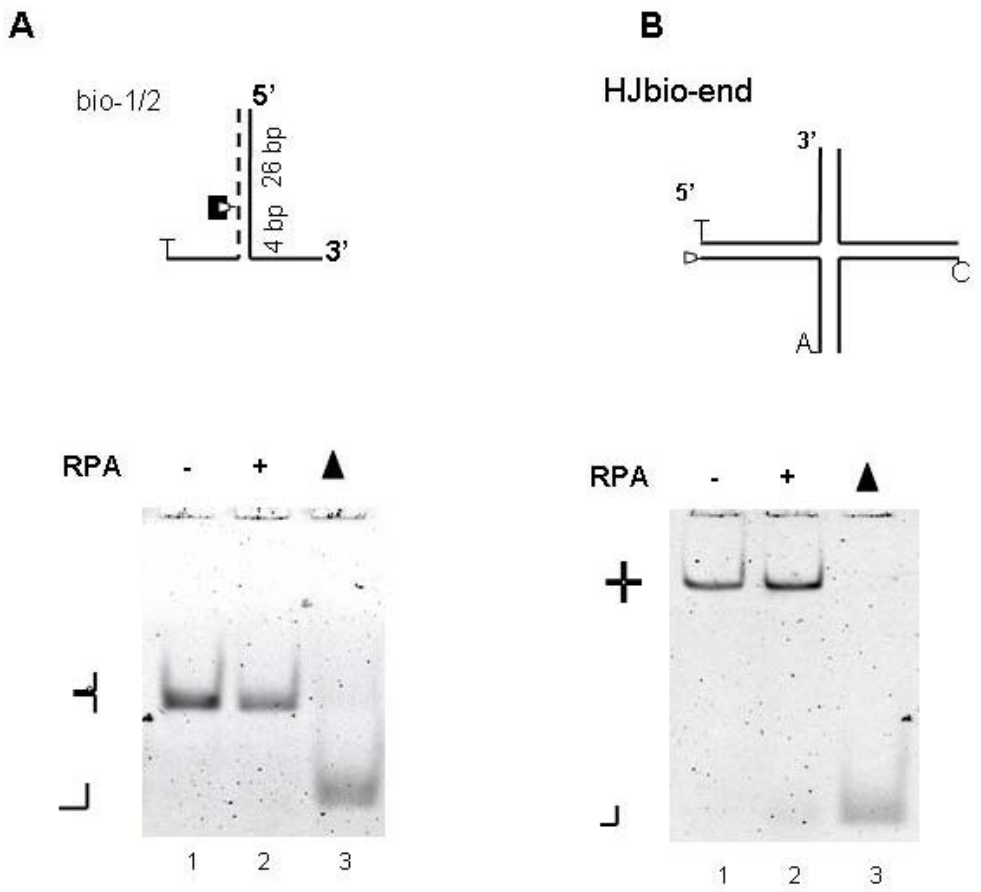

Fig. 3.16 RPA does not displace forked duplex or HJ substrates. (**A**) Reactions contained a 31-bp forked duplex (0.5 nM) with a biotinylated nucleotide (inverted trapezoid) on the nontranslocating strand. The forks were constructed by annealing oligonucleotides bio-1 and 2 (Table 1). T = TAMRA. The duplex length (bp) on either side of the biotinylated nucleotide is shown. The substrate was pre-incubated 0 (lanes 1) or 45 nM RPA (lane 2). Reactions were for 60 min at 37^{o}C, and were run on an 8% native polyacrylamide gel for 1.5 hrs and visualized with a Typhoon imager. The gel scans show the TAMRA emission channel**.** Boiled substrate (lane 3). **(B)** The HJbio-end substrate has a biotin tag at the 3' end of the A-oligo. T = TAMRA; C = Cy5; A = Alexa488; inverted trapezoid denotes a biotin moiety. The substrate was pre-incubated with 0 (lanes 1) or 45 nM RPA (lane 2). Boiled substrate (lane 3). The gel scans show the Cy5 emission channel**.** Reactions were for 60 min at 37^{o}C, and were run on an 8% native polyacrylamide gel for 1.5 hrs and visualized with a Typhoon imager

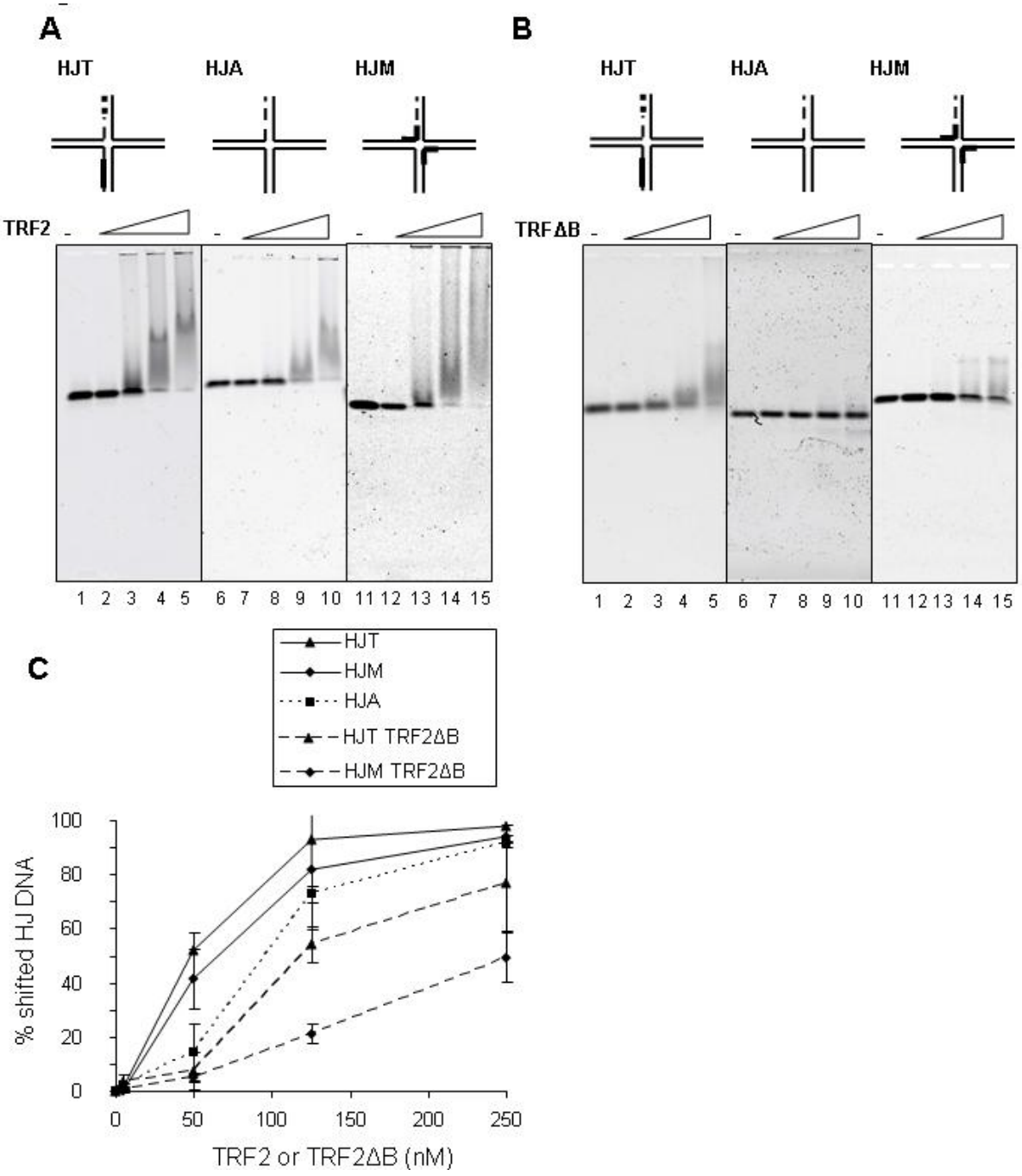

Fig. 3.17 TRF2 and TRF2ΔB exhibit increased binding to telomeric HJ constructs. The substrates (5 nM) telomeric arm HJT (lanes 1-5), non-telomeric HJA (lanes 6-10), and telomeric core HJM (lanes 11-15) were reacted with either 0, 5, 50, 125 or 250 nM TRF2 (**A**) or TRF2ΔB (**B**), respectively, in standard reaction buffer supplemented with 5% glycerol for 20 min at 24º C.

The reactions were loaded on a 1% 0.5X TBE agarose gel and electrophoresed for 1 hour at 140 V in 4ºC. Gels were scanned on a Typhoon Imager and visualized in the Cy5 channel. (**C**) Quantitation of the binding reactions. The percent of bound substrate was calculated as a function of total DNA in the Cy5 channel and plotted against TRF2 or TRF2ΔB concentration. Values represent the means and standard deviations from 2 or 3 independent reactions.

3.4 DISCUSSION

The WRN helicase/exonuclease protein is a RecQ family helicase that displaces HJ DNA, consistent with a cellular function in restoring stalled replication forks and disrupting homologous recombination intermediates (7;11;44;133). Previous studies showed that TRF2 protects telomeric HJ DNA at the t-loop/D-loop from cleavage by endonucleases *in vivo* (37;83) and *in vitro* (84), but TRF2 protection from HJ displacement was not previously reported. Our results showed that TRF2 protects HJ DNA from WRN-mediated strand displacement, but only if TRF2 can simultaneously bind the HJ with both of its DNA-binding domains: 1) the B domain which binds HJ DNA regardless of sequence, and 2) the Myb domain which binds telomeric duplex DNA (55;56). This is in contrast to previous research which showed that the B-domain is sufficient for protecting HJ DNA from enzymatic cleavage (84).

3.4.1 WRN mechanism of HJ strand displacement

The novel multiplex imaging of HJ strands in this study allowed us to elucidate mechanistic details of WRN strand displacement that would not be possible by labeling one strand of the construct as done previously. Our data are consistent with a mechanism by which WRN displaces HJs starting from the HJ center and processing outward. Specifically, on our substrate we propose WRN translocates 3' to 5' along the T-oligo from the center toward the 5' TAMRA label at the blunt end (**Fig. 3.6**). The first evidence was that a biotin-streptavidin barrier placed on the T-oligo at the HJ core led to a two-fold reduction in single stranded A- and T-oligo products (**Fig. 3.6**). Secondly, the barrier at the HJ core also induced some trapping of the A- and T-oligos into novel 2-way and 3-way products that lacked the C-oligo (**Fig. 3.7**) after WRN unwinding. This suggests that the biotin-streptavidin barrier inhibits WRN processivity along the A-T arm (**Fig. 3.6E,**I) and the T-unlabelled arm (Fig. 3.6E,III) depending on the barrier

location in the specific mobile core conformation, but not along the C-oligo arms which never contains the barrier in any core conformation. Thirdly, if WRN helicase initiated unwinding from an HJ arm blunt end and moved inward, then a barrier on the 3' end of a strand should block WRN helicase. However, such a barrier had no effect on strand displacement, nor did it lead to novel "trapped" products (**Fig. 3.8**). In summary, although the WRN exonuclease can initiate at an HJ blunt end (**Fig. 3.2**), the helicase does not, and thus, could not efficiently circumvent a protein bound to the HJ core such as TRF2.

Our results with the HJ barrier containing constructs corroborate electron micrographs showing WRN bound to the HJ center (142) and a protein-DNA cocrystal of RecG bound to a model replication fork (149). RecG and WRN belong to the SF2 superfamily of helicases (150), and RecG converts HJs into forked substrates and then into ssDNA similar to WRN (151)(11) (Fig. 1). The crystal structure shows a RecG monomer bound to the junction of a model replication fork (149), which resembles the junction of a regressed replication fork or HJ.

3.4.2 Mechanism of TRF2 inhibition of WRN telomeric arm HJ strand displacement

Our data help resolve a paradox of TRF2-WRN interactions. WRN is required for telomere preservation (18;152) and TRF2 stimulates WRN helicase on forked substrates (28), yet TRF2 prevents WRN promotion of aberrant telomere t-loop/D-loop HR cleavage (83). The WRN requirement for telomeric circle formation when the TRF2ΔB mutant is overexpressed (83) suggests TRF2 inhibits WRN-mediated branch migration at the telomere t-loop/D-loop. Here we report that TRF2 modulation of WRN activity is highly substrate dependent. At a 50:1 TRF2:HJ molar ratio, prebound TRF2 inhibited WRN strand displacement of an HJ with two telomeric arms by 50%, but unwinding of the non-telomeric HJ was unaffected (**Fig. 3.7**). At this same ratio nearly 100% of both the telomeric arm and non-telomeric HJs are bound by TRF2 (**Fig.**

3.17), consistent with previous studies (56). This shows that TRF2 substrate binding via the B domain only (non-telomeric HJ) provides no protection against WRN activity (**Fig. 3.7B**). Consistent with this, the TRF2 Myb domain alone cannot protect the telomeric arm HJ (**Fig. 3.7C**) or a telomeric fork (28) (the B-domain does not bind a fork), from WRN activity.

The mechanism of TRF2 protection against HJ cleavage differs from protection against WRN HJ strand displacement. A previous study showed that a 40:1 TRF2:HJ ratio reduced enzymatic HJ cleavage by 80% on an HJ with two telomeric repeats in the mobile center, and that protection was primarily via the B domain (84). With this same substrate (HJM, **Fig. 3.9A**) even a 50:1 molar ratio of TRF2 to HJ provided no protection against WRN helicase activity (**Fig. 3.9B**). One explanation for this difference is that while this HJ has binding sites for both the TRF2 B and Myb domains, both probably cannot bind at once. The footprint of a single Myb domain is YTAGGGTTR (153), thus, a TRF2 homodimer cannot bind this substrate with both Myb domains at once. A single Myb domain can bind only if the HJ mobile core has migrate to one extreme (**Fig. 3.9A**), but it may not be sterically possible for the B domain to bind as well. Consistent with this, when the HJ-specific TRF2 B-domain was deleted the percent of TRF2 bound telomeric core HJM was half that of telomeric arm HJT (**Fig. 3.12C**). Protection against WRN helicase activity requires both TRF2 domains to be bound, not just the B-domain as is the case with protection against cleavage. Poulet and colleagues reported that the TRF2 B-domain "melts" the mobile center of the HJ, creating an altered conformation that HJ cleaving enzymes cannot recognize (84). A melted HJ center would hardly discourage a helicase, particularly one that initiates displacement at the HJ center, such as WRN (**Figs. 3.5-3.6**).

Several potential models might explain the TRF2 inhibition of WRN activity on telomeric arm HJ. In stark contrast to a biotin-streptavidin complex at the HJ core (**Fig. 3.5**), a

core bound TRF2 B-domain does not inhibit WRN (**Fig. 3.11**), which rules out a TRF2 steric inhibition model. Since TRF2 interacts with WRN via the B-domain (83), we would have expected inhibition on a non-telomeric HJ, which was not the case (Fig. 3.7B). Thus, TRF2 inhibition of WRN telomeric HJ activity is unlikely to be mediated by protein-protein interactions, which is the proposed mechanism for p53 inhibition of WRN HJ activity (154). Instead, we favor a model whereby both TRF2 DNA binding domains must bind the HJ simultaneously in order to "lock down" the substrate into a conformation that is unfavorable for WRN strand displacement or branch migration activity. While HJ constructs are often depicted as square-planar shapes, crystallographic evidence indicates that duplex arms can stack on one other (155). Single-pair FRET and structural studies showed that HJ DNA alternates between the stacked conformation and an open square-planar conformation, which is the form that branch migrates and is favored by proteins that promote HJ migration (149;156-158). Simultaneous interaction of TRF2 with the core and the telomeric arms might promote the stacked arm conformation, perhaps via TRF2 molecules bound to separate arms interacting with each other through dimerization or higher oligomeric forms. Consistent with this, the telomeric arm HJ (HJT) exhibited a slower TRF2-bound migrating species than either the non-telomeric (HJA) or telomeric core HJ (HJM) (**Fig. 3.12A**) which likely reflects a larger oligomeric complex. Our results agree with previous studies that show TRF2 has higher affinity for HJs with telomeric repeats, compared to non-telomeric HJs (84). Therefore the number of accessible telomeric binding sites is important for TRF2 HJ binding, and this correlates with protection from WRN activity. We suggest that a fully telomeric HJ would enjoy greater TRF2 protection since there are far more TRF2 binding sites compared to our hybrid constructs.

3.4.3 Role for RAP1/TRF2 complex in protection against HJ displacement

The current study extends our understanding WRN function in the context of the mechanism of action of shelterin proteins at telomeres. While RAP1 forms a complex with TRF2 and is necessary to protect telomere ends from NHEJ (58), it does not alter TRF2-mediated inhibition of WRN on an HJ containing telomeric arms (Fig. 3.9). RAP1 binds to the central TRF2 dimerization domain (148), whereas WRN interacts with the TRF2 N-terminal B domain (83). Thus, our biochemical data is consistent with cellular studies since RAP1 should still localize to telomeres with TRF2ΔB expression but offers no protection against aberrant HJ cleavage (83), and RPA1 does not localize to telomeres in the absence of TRF2 when protection against aberrant NHEJ is lost (159). Neither does RAP1 increase the binding affinity of TRF2 for telomeric DNA (160). Therefore, RAP1-mediated protection against NHEJ probably has no mechanistic basis in protection against WRN activity.

3.4.4 The Yin and Yang of WRN at telomeres

Telomeres are dynamic structures regulated by the interplay of multiple proteins composing both the shelterin complex and the numerous proteins that interact with shelterin. Many of these proteins become deleterious for telomere maintenance if dysregulated (1;161). WRN is necessary to prevent telomeric circle formation in telomerase-positive cells, but without full-length TRF2, WRN itself is required for telomere circle formation (45). In yeast, the TRF2 ortholog Taz1 similarly protects against telomere loss mediated by the WRN ortholog Rqh1 (162) in conjunction with the RPA ortholog Rad11 (163). Telomere preservation is restored in Taz1 mutants by either mutating a SUMOylation site on the Rgh1 protein (162) or by overexpressing the shelterin protein POT1 (163). By demonstrating that TRF2 can protect

telomeric HJ DNA from WRN activity, our study provides a mechanistic basis for the cellular studies showing that TRF2 or Taz1 regulate WRN or Rqh1 activity, respectively.

In summary, we verified that WRN displaces HJs from the center and progresses outward, and that this strand displacement occurs simultaneously with WRN exonuclease activity. We found that TRF2 can protect telomeric HJs from WRN activity only if both TRF2 DNA binding domains simultaneously engage the HJ core and the arms. Our biochemical studies provide an explanation for the cellular studies that indicate WRN adversely affects telomere processing in the absence of TRF2 (83). Understanding how shelterin accessory factors, such as WRN, are regulated by specific shelterin proteins is critical for defining mechanism of telomere loss and preservation, and the consequences of dysregulation for human disease.

FUNDING

We acknowledge Dr Walter Chazin RPA protein, the production of which was supported by National Institutes of Health grant GM65484 (W.J.C.). This research was supported by the Ellison Medical Foundation (PLO), the National Institute of Environmental Health Sciences grant number ES0515052 (PLO) and the National Institute on Aging grant number F30AG032861 (GJN).

ACKNOWLEDGEMENTS

We thank the members of the Opresko lab for their critical reading of the manuscript and feedback.

4.0 GENERAL DISCUSSION

The shelterin complex and telomere-associated proteins must modulate non-canonical DNA structures in order to maintain telomeric and genomic stability. While telomere length has captured the public's attention as a cause for aging, cellular senescence may occur as a result of a single telomere undergoing sudden shortening, thereby triggering an unwanted DNA damage response. This study sought to elucidate the DNA structures and mechanisms underlying the known cellular mutational and knockout models for the shelterin and telomere-associated proteins. Our goals were to: (1), find how G4 DNA is arranged on long, physiologically relevant telomeric substrates, (2), elucidate how POT1 competes with G4 folding on long telomeric substrates, (3), determine the mechanism of WRN displacement of HJ DNA, and (4), determine whether TRF2 protects HJ DNA from WRN displacement.

4.1 G4 DNA IS UNDERFOLDED AND FORMS BEADS ON A STRING ON LONG TELOMERIC ssDNA

The literature on G4 DNA formation is unclear on how G4 DNA is arranged on long, physiologically relevant telomeric ssDNA. One CD study implied that G4 DNA folds to maximize the number of quadruplexes that can be folded on a substrate (78), creating a beads on a string arrangement of individual quadruplexes separated by linker ssDNA (**Fig. 2.1**). In contrast another group found that G4 DNA formation (as measured by CD) was enthalpically penalized by progressively longer telomeric substrates (77), implying that even if the quadruplexes formed beads-on-a-string, they might not fold into the maximum amount of G4 DNA (79).

We have used AFM to visualize and quantify G4 DNA on long telomeric ssDNA. We used the length of the G4 DNA (defined as being over 1 nm in height) and demonstrated that

$(TTAGGG)_{16}$ molecules on average have twice the length of G4 DNA as $(TTAGGG)_8$ and $(TTAGGG)_4$ molecules (**Fig.2.2D**). We confirmed that under our deposition conditions that ssDNA and duplex DNA controls did not form structures over 1 nm. We investigated why $(TTAGGG)_{16}$ was underfolded by a factor of two with a simple combinatoric model, showing that as an ensemble of all possible conformations, that 2 quadruplexes were the most probable. However, our model predicted a nearly equal quantity of triple-quadruplex forming molecules, which was not the case. We found that our folding profile matched folding curves for ssDNA composed of 13 or 14 repeats better than 16 or 15.

A recent study used exonuclease degradation to show that when the 5' end of a telomeric substrate is bound to a flanking sequence, G4 DNA formation is favored at the 3' end, with the probability of G4 DNA formation decreasing as one moves inward (164). We therefore hypothesize our 34 bp stem on the 5' end of the telomeric ssDNA inhibits G4 folding on the neighboring 2-3 repeats, which is why the folding profile of the G4 DNA resembles a 13- or 14-repeat substrate more than one with 16 or even 15 repeats.

Due to the limitations of the AFM resolution (see below), we measured overall G4 DNA length, but were still able to visualize substrates with 2-3 peaks (**Fig. 2.3**), supporting a beads-on–a-string structure of individual quadruplexes linked by ssDNA. Had the guanine tetrads aligned themselves into a more rigid array, we would have expected a taller (i.e., less deformable by the AFM probe) structure with no peaks or surface features like the G-wire.

One may define the conditions that permit our system to resolve individual peaks. Bustamante and Keller defined an equation (cited in 88) relating the minimum resolution *d* to the radius of the AFM tip and the "dimple depth" between objects (which is 0.2 nm in the case of defining G4 peaks).

$$d = \sqrt{2R}(\sqrt{\Delta z} + \sqrt{\Delta z + \Delta h})$$

where d is the resolving distance, z is the "dimple", h is the difference of the height between the peaks and R is the radius of the tip. This assumes the case where the resolution *d* is greater than $\sqrt{2R\Delta h}$. Using the manufacturer's (Nanosensors^{TM}) specifications, the tip radius is 7 nm, and assuming the peaks are the same height (Δh=0), the resolution is 3.6 nm. The G4 DNA structure in **Fig.1.5** is 4.4 nm along its longest axis, which puts G4 DNA at the edge of the system's capabilities even when the tip is not damaged or blunted.

We used Pymol to make two copies of a G4 DNA structure (Protein Data Bank ID: 2JPZ) and bond the two copies together (**Fig. 1.5**), providing a model of what a double-quadruplex beads on a string structure would look like. Using the distance function on Pymol, the distance between the midpoints of the two quadruplexes is about 2.2 nm, below the resolution of the ideal calculated above (3.6 nm) by 1.4 nm. Following Hansma and colleagues(107), ssDNA has a length of 0.2-0.5 nm per nucleotide, depending the degree of extension or nonspecific folding. Therefore a resolution of 3.6 nm requires a gap equivalent of 3-7 nucleotides in addition to the 3 nucleotide (TTA) linker already present between the quadruplexes. When we resolve two peaks, there must be at least 1.5 telomeric repeats present between the quadruplexes, though there may be more since the ssDNA may not be fully extended. Since the mean interpeak distance of G4 peaks on Tel16 was 20 nm, this translates to roughly 7 telomeric repeats, assuming that they are fully extended. Since a maximally folded $(TTAGGG)_{16}$ has only TTA linkers and no "spare" telomeric repeats to use as spacer, one would not expect to see four distinct peaks under the above criteria for determining peaks.

4.1.1 POT1 is a steric driver on long telomeric tails

POT1 was shown to passively compete with G4 folding on short telomeric substrates by binding to spontaneously unfolded telomeric ssDNA and thus preventing the DNA from refolding(103). POT1, however, was never reported on longer, more physiologically relevant telomeric substrates, and until our own results the distribution and structure of the G4 DNA on the longer telomeric substrates were themselves mysteries.

We found that POT1 competes with and destabilizes G4 DNA, with the average length of G4 DNA being 12.5 nm, half of the G4 DNA formed by $(TTAGGG)_{16}$ alone and with the population of G4 DNA reduced to 22% of the molecules recorded. Interestingly, POT1 binds $(TTAGGG)_{16}$ and the non-G4 forming control substrate in a roughly equal distribution of lengths, as depicted in **Fig. 2.10E**. This suggests that POT1 can destabilize G4 DNA, what we dubbed a "steric driver", since otherwise POT1 would be limited to unfolded parts of the $(TTAGGG)_{16}$ molecule, while in the control substrate the whole length of the ssDNA is available for binding.

The steric driver effect may in fact be an emergent cooperative effect, as a bound POT1 would destabilize existing G4 DNA and thus open more ssDNA to POT1 binding, but one would have to test a range of POT1:DNA ratios to determine this. One limiting factor in AFM is that if one "carpets" the mica with too much protein, it becomes hard to determine whether individual POT1 molecules are on the same substrate or not.

4.1.2 Future directions

The WRN helicase/exonuclease unwinds G4 DNA and localizes to the telomere during S-phase. WRN is currently not at the level of purity necessary for AFM, since there are breakdown products contaminating the full-length WRN. However, a "double-tag" vector for WRN,

employing both hexahisitidine and FLAG affinity tags instead of just hexahistidine(144), promises to increase the purity of WRN preparations. An AFM-suitable WRN sample would provide an important structural underpinning to our biochemical characterization of WRN action on HJ DNA (**4.3**) with the added advantage that AFM provides a 3D image without the use of heavy metal contrast agents, such as with a recent electron microscopy (EM) study of WRN binding to HJ DNA (142). Another route to study WRN on G4 DNA is Förster Resonance Energy Transfer, discussed in Chaper **4.2**, and unlike AFM is better optimized for mechanistic, time-dependent details of helicase activity.

POT1 competes with RPA binding on telomeric, G-rich ssDNA, where RPA can act as a DNA damge reporter while POT1 suppresses ATR (165). Since POT1 has the greater affinity for telomeric ssDNA than RPA, there is a higher POT1:RPA ratio at telomeres in wild type cells(42), but the knocking down of POT1 expression would increase the signal to trip the repair machinery. A structural study using AFM could demonstrate how the G4 DNA landscape in $(TTAGGG)_{16}$ is altered by the more physiologically relevant combination of RPA and POT1, rather than just POT1. We already have a physiologically relevant model of the telomeric 3' tail in our $(TTAGGG)_{16}$ construct, but would need a means to discriminate between the two proteins. Quantum dots and gold nanoparticles conjugation to the proteins would provide one such option.

Quantum dots (Qdots) have been conjugated via two antibodies to UvrB and used to demonstrate UvrAB complex formation on nicked DNA(111). While the Qdots were successful in preserving the binding functionality of UvrAB, their large size (radius ~20 nm)(111) might create artifacts with RPA-POT1 coating of a DNA strand. A recent paper on thiol-functionalized 2 nm gold particles might provide a better route for studying ssDNA-binding proteins on the Tel16 substrate(166).

A fringe benefit of the gold nanoparticles is that their regular shape could be of service for deconvoluting the original shapes of the molecules scanned by the AFM. The image viewed on the computer screen shows the convolution of the shape of the protein plus that of the tip shape. The tip shape is usually a matter of conjecture unless one takes the trouble to put each tip under a scanning electron microscope to record its shape. Even then, as soon as the tip comes into use it will be blunted. By having a gold standard in the sample with a well-defined shape, it is possible to back-calculate the shape of the tip, and thus, one could be able to properly deconvolute the shapes of the other non-gold molecules in the image.

4.2 TRF2 PROTECTS HJ DNA FROM WRN ACTIVITY IF BOTH DOMAINS ARE BOUND

The shelterin protein TRF2 plays a multifaceted role in protecting the telomere. It suppresses the non homologous end joining (NHEJ) of double strand break repair at the telomere by localizing the NHEJ-suppressor RAP1 to the telomere(83), and is also known to suppress sudden telomere loss mediated by Holliday Junction cleaving enzymes. The latter phenomenon is dependent on WRN, as cells with knocked down WRN protein do not undergo telomere loss when TRFΔB is expressed in the cell(83). We verified that TRF2 inhibits WRN-mediated displacement of HJ DNA if TRF2 could simultaneously bind the HJ with both DNA binding domains. Using a novel fluorophore-based assay, we also demonstrated that WRN helicase initiates strand displacement in the center of the HJ and moves outward.

While the B domain of TRF2 primarily conferred HJ protection against GEN1-mediated cleavage(84), the B domain alone cannot protect HJ DNA from WRN activity, as evidenced by the HJA substrate, which can only be bound via the B-domain and is totally unprotected by

TRF2 (**Fig. 3.12B**). Similarly the Myb domain alone conferred no protection to HJT, which has Myb binding sites in two arms (**Fig. 3.12C**). While the HJM substrate, with two telomeric repeats in the center, was protected by TRF2 from GEN1, it was not protected with WRN, either with or without exonuclease (**Fig. 3.14**). Since TRF2 can bind HJM with only one domain at a time, and only one TRF2 monomer can bind the telomeric repeats at one of the extremes of branch migration (Fig. A), the HJM cannot be bound simultaneously by both TRF2 domains. The simultaneous binding of HJT by both TRF2 domains (with dimers being able to bind both arms) is suggested in the EMSA results, where HJT is shifted to a slower migrating band than the other HJ substrates studied (**Fig. 3.17**).

We also elucidated the mechanism of WRN-mediated strand displacement by means of biotin-streptavidin barriers one of the oligonucleotides making up the Holliday Junction. A centrally placed biotin-streptavidin barrier trapped two fluorophore-labeled oligonucleotides as forked and three way species, with diminished liberation of ssDNA species (**Fig. 3. 6**). In contrast, a biotin-streptavidin barrier on the 3' end of one oligonucleotide (where WRN might load if it initiated unwinding from the edge) had no effect on HJ DNA displacement (**Fig. 3.7**).

4.2.1 Future directions for WRN and telomeric HJ DNA

We proposed a mechanism for TRF2-mediated HJ protection by altering the geometry such that the HJ is not readily branch-migrated. Single molecule FRET is a method that would readily test this model. In a proposed model HJ with telomeric-repeat containing arms in *trans* arrangement (**Fig. 4.1**), the HJ can be either in a square planar (I) or stacked (II) conformaitons. In (I), a donor dye (green) remains distant from the acceptor (red), so the latter does not emit. However, if the *trans* arms are brought close together (II), such as via TRF2 binding and multimer coordination, the acceptor will emit. Care must be taken to avoid the "quenching" effects of G-rich DNA with

fluorophores, where the excitation energy of the dye is transferred to the guanine(167). Fig. 4.2 depicts one such precaution by placing the fluorophores on the C-rich arms instead of the G-rich. The ends of the telomere-containing arms may also have a brief non-telomeric T/A sequence right at the edge in order to protect the signal.

The HJ constructs in this study used “fluid” or homologous centers permitting branch migration. Researchers have already used FRET to study branch migration and HJ arm stacking simultaneously(149;156-158). Thus one may monitor both the effect of TRF2 on HJ geometry and WRN on branch migration. Single molecule FRET would provide a powerful model system for elucidating the mechanism TRF2 at telomeric HJ DNA, and also further exploring the mechanism of WRN-mediated branch migration of HJ substrates.

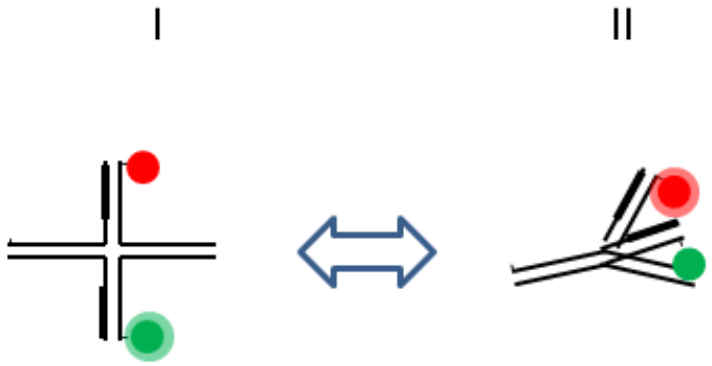

Fig. 4.1 Schematic of FRET-based assay for HJ geometry and branch migration. An HJ construct with two telomeric-repeat (heavy black denotes G-rich strand) arms in *trans* has donor (green) and acceptor (red) fluorophores on the opposing arms. A square planar, branch-migrating conformation (II) will have a low FRET output. When HJ arms associate with the arm across from them in a stacked confirugation, FRET efficiency will be high (II). Biotin tags to link the construct to the flow cell surface have been omitted for clarity.

4.3 SINGLE PAIR FRET ANALYSIS OF G4 DNA FORMATION AND MODULATION

Our studies of G4 DNA focused on a population of quadruplexes on long telomeric tails. In addition to POT1, however, the RecQ helicases BLM and WRN as well as the RecD helicase FANCJ actively unwind G4 DNA. Biochemical studies have tested these helicases against intermolecular quadruplexes formed from two or more oligonucleotides, with one of the oligonucleotides radiolabelled(30;128). The G4 DNA most likely encountered in human telomeres is intramolecular, however, such intramolecular quadruplexes defy bulk biochemical studies since they can readily "snap back" into G4 DNA after being unwound by a helicase. Single pair Forster Resonance Energy Transfer (spFRET) utilizes a non-radiative transfer of energy from a donor dye to an acceptor, the transfer of energy is inversely proportional to the sixth power of the distance between the dyes, providing a real-time "molecular ruler" to track folding of G4 DNA. Single-pair FRET has been used to study G4 DNA folding at the single-molecule level, however it was done on a construct where the telomeric DNA was in a reversed orientation, with the telomeric ssDNA being $GGG(ATTGGG)_3$ rather than $(TTAGGG)_3GGG$, perhaps as a result of the contemporary limitations on synthesizing an oligonucleotide with a biotin at one end(123). Furthermore, the authors did not examine G4 DNA folding in the context of telomere-associated proteins like POT1 or WRN(123). We constructed a FRET substrate that has the telomeric ssDNA in the correct orientation. As with the previous studies, we have a duplex stem, with the complementary strand having the donor dye (TAMRA), and the 5' end of the telomeric repeat-containing oligonucleotide having a biotin tag, used to fix the construct to the surface of the flow

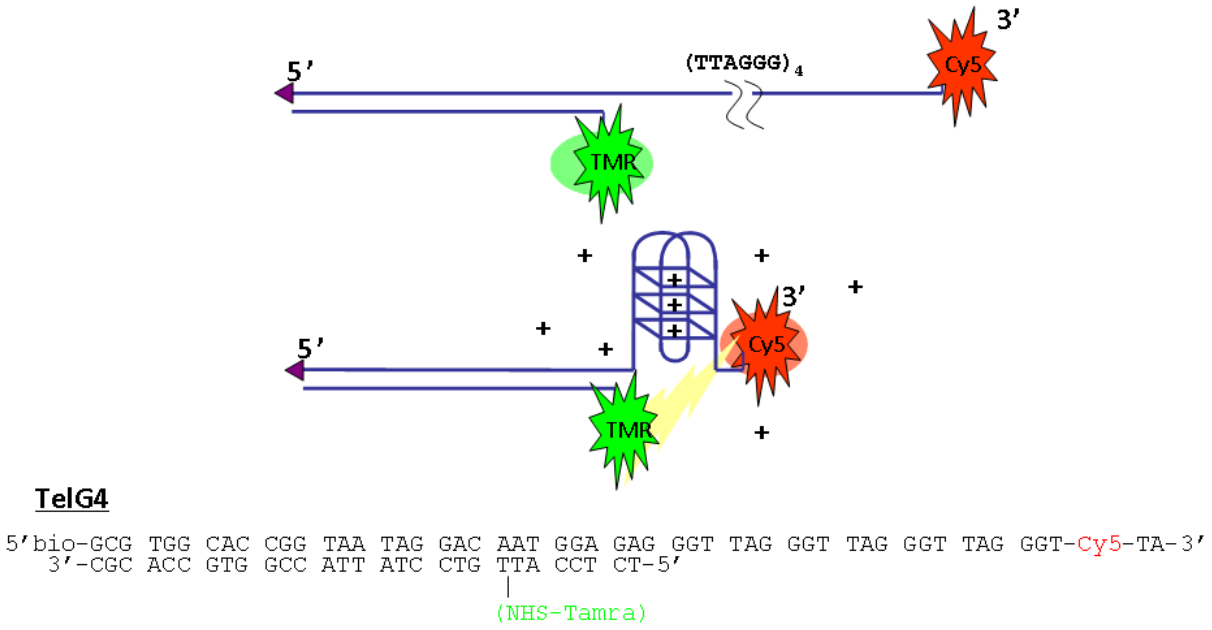

Fig. 4.2 Schematic of the TelG4 oligonucleotide in both its unfolded (above), folded (middle) conformations. The sequences and placement of fluorophores and biotin are indicated at bottom. Triangle denotes biotin in upper and middle schematics. "+" signs denote cations (e.g., K^+) which are necessary for folding and stabilization of the individual tetrads. Courtesy of Matthew Fagerburg.

cell described below. The acceptor dye (Cy5) is on the end of the telomeric repeats, but is not covalently linked to the terminal guanines, which can quench the signal of the acceptor dye.

Total Internal Refraction FRET (TIRF) uses an inverted microscope and a flow cell created from a plastic cover and a quartz slide with a thin spacer in between. The flow cell is illuminated above by evanescent light from a prism that is illuminated by total internal refraction (TIR) from a laser source (laser wavelength is 532 nm for our apparatus). The evanescent light "leaking" from the TIR at the prism boundary excites fluorophores in the flow cell without creating excessive background illumination. The flow cell surface is covered with biotinylated bovine serum albumin (BSA) and streptavidin, to which our construct (which has a biotin tag) binds. The fluorescent emissions are collected by the objective and split into two channels by a dichroic mirror, one each for the donor and acceptor emission spectra, which are quantitated and correlated with each other.

4.3.1 Results

We have verified that our oligonucleotide transitions from a low FRET state with no potassium ions to a high FRET state in the presence of potassium (100 mM), corresponding to the folded quadruplex. We incubated 50 μL of 3 pM FretG4 in standard reaction buffer (40 mM Tris, pH 7.5, 1% beta-mercaptoethanol, 0.8 % glucose, 0.1 mg/mL glucose oxidase, and 0.02 mg/mL catalase) in a reaction chamber and flushed the reaction volume (50 μL) after 5 minutes with standard reaction/scavenger buffer before imaging. Data were collated at 30 millisecond intervals for 90 seconds. Intensities over a 1 second period were averaged and the resulting normalized FRET efficiencies ($E_{acceptor}/E_{donor}$) were calculated on MatLab (Mathworks, Natick, MA) and

plotted as histograms on Excel (Fig. 4.4). We found that a high FRET efficiency (0.8) was achieved when KCl was present, consistent with G4 DNA folding (Fig. 4.4A,B).

To test whether POT1 inhibited G4 folding, we prepared a sample with standard reaction buffer, 100 mM KCl and incubated 20 nM POT1 with 3 pM FretG4 substrate for 20 minutes at room temperature before injecting into a flow cell and visualizing. We found that the unfolded peak in the histogram reappeared (Fig. 4.4C), indicating that there was small population of unfolded DNA under conditions where all DNA was folded initially. The relatively small response from such a massive quantity of POT1 is perhaps attributable to the adsorption of POT1 to the BSA coated surface.

We verified that POT1 bound TelG4 under bulk conditions using a Cary Eclipse fluorimeter (Varian, Palo Alto, CA). Following ionic concentrations in which POT1 binding is known to occur(103), we incubated 10 nM FretG4 DNA in 40 mM Tris, pH 7.5, 1 mM $MgCl_2$, 15 mM NaCl, 50 mM KCl, 1% beta-mercaptoethanol, and 5% glycerol with or without 30 nM POT1 for 30 minutes. We found a reduction in the 587 nm emission peak, signifying less G4 DNA folding and thus a lower FRET efficiency with a far lower POT1:DNA ratio than in the single molecule studies. Further work is necessary to determine whether the salts, the temperature or the surface conditions of the spFRET led to the effectively lower POT1 affinity for G4 DNA in the single molecule FRET relative to the bulk study. The use of polyethylene glycol (PEG) preparations instead of BSA to coat flow chambers may also prevent protein adsorption, but we must figure out the proper PEG formulation and coating protocol.

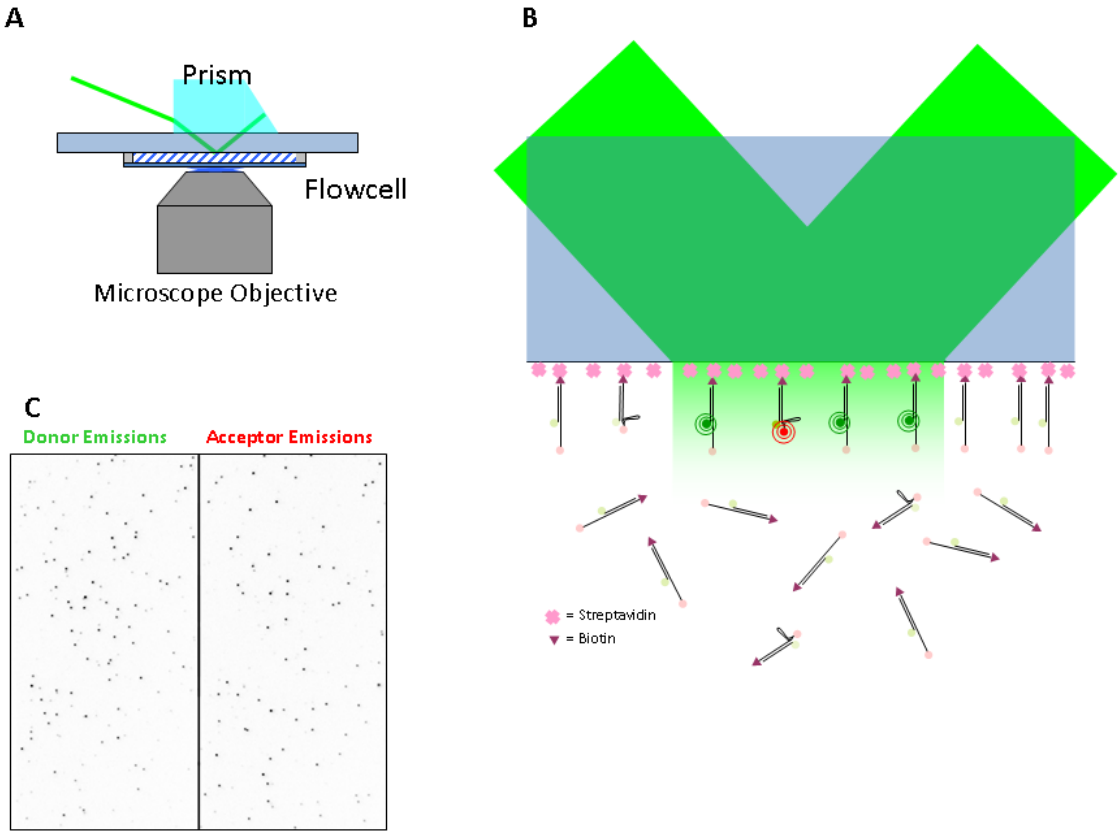

Fig. 4.3 Schematic of single molecule, wide field FRET microscopy of G4 DNA. **A**, Schematic of the TIRF assembly on an inverted microscope, showing the prism, laser source, flow cell and objective. **B**, Schematic of laser light refracting off the quart slide, with an evascent field exciting dye-labelled constructs bound to the surface underneath. **C**, Example data showing the microscopic wide field images with the acceptor and donor emission channels from the same part of a slide. Courtesy of Matthew Fagerburg.

In order to elucidate the mechanism of WRN unwinding of G4 DNA, we had to verify that WRN was active under FRET conditions. We have verified that WRN is active under modified FRET conditions at room temperature. WRN was titrated in the concentrations with 0.5 nM forked substrate (Tel3/4) with 22 bp of duplex DNA as well as a 3' and 5' tails which was radiolabelled and annealed as described in (33). WRN was reacted at room temperature in 40 mM Tris pH 7.5, 4 mM $MgCl_2$, 0.1 mg/mL BSA, 0.26% glucose, 2 mM ATP, 0.067 mg/mL glucose oxidase, 0.013 mg/mL catalase, and 100 mM KCl. We titrated WRN against beta-mercaptoethanol (BME) concentrations as noted in (**Fig. 4.6**). BME is a free radical scavenger that protects against dye-blinking, but can also denature proteins and degrade activity. DTT was omitted to avoid noisy FRET imagery, probably due to DTT precipitates in the solution. We found that WRN activity against Tel3/4 was increased with greater quantities of BME, probably because it protected WRN against oxidation in the absence of DTT. Therefore WRN is likely to be active under FRET conditions.

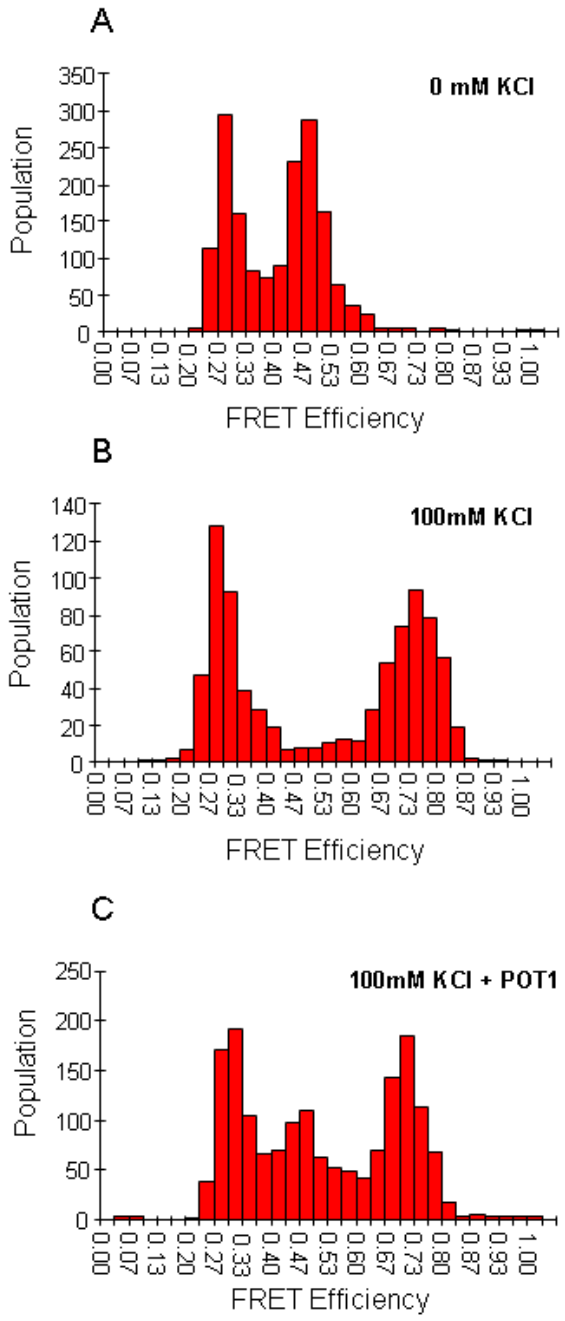

Fig.4.4 G4 DNA folding requires KCl but is partially inhibited by POT1. 3 pM FretG4 is reacted under standard conditions alone (**A**), with 100 mM KCl (**B**) or with 100 mM KCl and 20 nM

POT1 (**C**) before quantitating by spFRET. Histograms were created as described in Materials and Results.

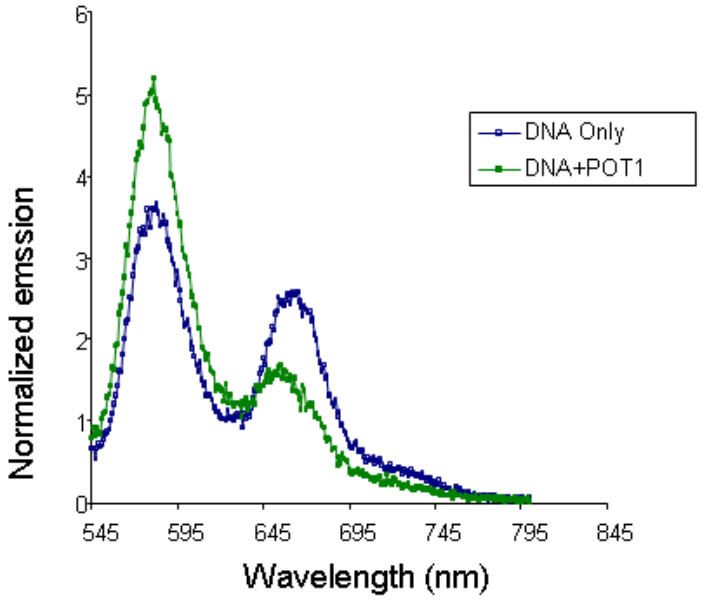

Fig. 4.5 POT1 reduces FRET signal attributable to G4 DNA folding. Fluorescence levels of G4 DNA in bulk fluormetry buffer alone (blue) or in addition to POT1 after 30 minutes of incubation at room temperature (green) are scanned in a Cary Eclipse fluorimeter.

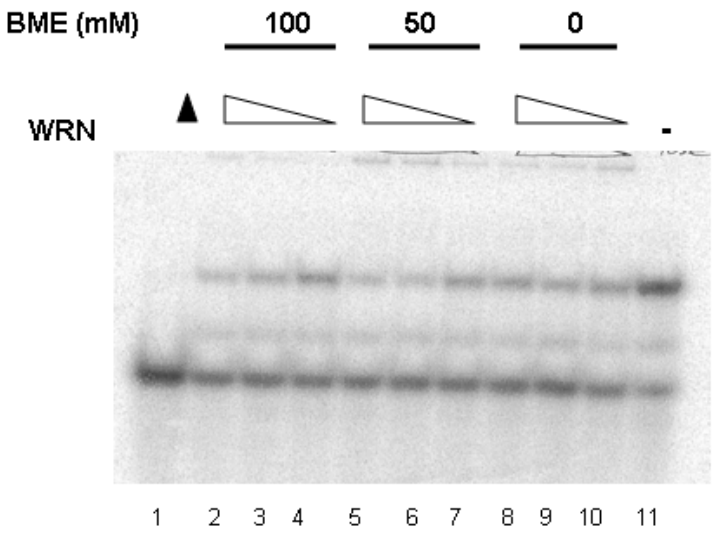

Fig. 4.6 WRN is active at room temperature with FRET conditions. 0.5 nM Tel3/4 substrate (0.5 nM) was incubated with 0.25 (lanes 4,7,10), 0.5 (lanes 3,6,9), or 1 nM (lanes 2,5,8) WRN incubated at room temperature for 30 min in 0 (lanes 2-4), 50 (lanes 5-7) , or 100 mM (lanes 8-10) BME . Reactions were stopped with 11.5 μL stop dye(33), and run on a 12% native polyacrylamide, 1X TBE, 0.1% SDS gel at 150 V for 90 minutes. Gel was exposed overnight and scanned on a GE Storm.

4.4 CONCLUSIONS

The human telomere is critical for maintaining the stability of genomic DNA. Even non-dividing cells must be protected from unwanted DNA damage responses, and failure to do so leads to cell senescence and tissue dysfunction. The telomere is composed of both repetititive DNA and the shelterin complex of proteins which cap the telomeric DNA. Researchers characterized numerous phenotypes for mutants and knockdowns of shelterin and telomere-associated proteins like WRN which interact with the shelterin proteins. While rich and provocative, the cellular data cannot resolve the mechanisms underlying the phenotypes. Non-canonical, dynamic DNA structures like Holliday Junctions and G-quadruplexes further complicate the task of elucidating mechanisms *in vitro.* Single molecule methods such as AFM and spFRET provide an invaluable means to characterize the heterogeneous, dynamic DNA structures present at the telomere. We have characterized important mechanism of DNA structural modulation in both the “open” (i.e., undergoing replication, elongation or repair) telomere and also the closed (i.e., cells outside of S-phase of non-dividing cells).

5.0 BIBLIOGRAPHY

1. Palm,W. and de Lange, T. (2008) How shelterin protects mammalian telomeres. *Annu. Rev. Genet.*, **42**, 301-334.

2. Harley,C.B., Futcher,A.B. and Greider,C.W. (1990) Telomeres shorten during ageing of human fibroblasts. *Nature.*, **345**, 458-460.

3. Blasco,M.A. (2007) Telomere length, stem cells and aging. *Nat. Chem. Biol.*, **3**, 640-649.

4. Chen,J. (2004) Senescence and functional failure in hematopoietic stem cells. *Exp. Hematol.*, **32**, 1025-1032.

5. Minamino,T., Miyauchi,H., Yoshida,T., Ishida,Y., Yoshida,H. and Komuro,I. (2002) Endothelial cell senescence in human atherosclerosis: role of telomere in endothelial dysfunction. *Circulation.*, **105**, 1541-1544.

6. Munoz,P., Blanco,R., Flores,J.M. and Blasco,M.A. (2005) XPF nuclease-dependent telomere loss and increased DNA damage in mice overexpressing TRF2 result in premature aging and cancer. *Nat. Genet.*, **37**, 1063-1071.

7. Opresko,P.L. (2008) Telomere ResQue and preservation--roles for the Werner syndrome protein and other RecQ helicases. *Mech. Ageing Dev.*, **129**, 79-90.

8. Flores,I. and Blasco,M.A. (2009) A p53-dependent response limits epidermal stem cell functionality and organismal size in mice with short telomeres. *PLoS. One.*, **4**, e4934.

9. Minamino,T., Orimo,M., Shimizu,I., Kunieda,T., Yokoyama,M., Ito,T., Nojima,A., Nabetani,A., Oike,Y., Matsubara,H. *et al.* (2009) A crucial role for adipose tissue p53 in the regulation of insulin resistance. *Nat Med.*, **15**, 1082-1087.

10. Ju,Z., Jiang,H., Jaworski,M., Rathinam,C., Gompf,A., Klein,C., Trumpp,A. and Lenhard,R.K. (2007) Telomere dysfunction induces environmental alterations limiting hematopoietic stem cell function and engraftment. *Nat. Med.*, **13**, 742-747.

11. Kudlow,B.A., Kennedy,B.K. and Monnat,R.J., Jr. (2007) Werner and Hutchinson-Gilford progeria syndromes: mechanistic basis of human progeroid diseases. *Nat. Rev. Mol. Cell Biol.*, **8**, 394-404.

12. Huang,S., Lee,L., Hanson,N.B., Lenaerts,C., Hoehn,H., Poot,M., Rubin,C.D., Chen,D.F., Yang,C.C., Juch,H. *et al.* (2006) The spectrum of WRN mutations in Werner syndrome patients. *Hum Mutat.*, **27**, 558-567.

13. Martin,G.M. (2005) Genetic modulation of senescent phenotypes in Homo sapiens. *Cell.*, **120**, 523-532.

14. Baird,D.M., Davis,T., Rowson,J., Jones,C.J. and Kipling,D. (2004) Normal telomere erosion rates at the single cell level in Werner syndrome fibroblast cells. *Hum Mol Genet.*, **13**, 1515-1524.

15. Schulz,V.P., Zakian,V.A., Ogburn,C.E., McKay,J., Jarzebowicz,A.A., Edland,S.D. and Martin,G.M. (1996) Accelerated loss of telomeric repeats may not explain accelerated replicative decline of Werner syndrome cells. *Hum Genet.*, **97**, 750-754.

16. Laud,P.R., Multani,A.S., Bailey,S.M., Wu,L., Ma,J., Kingsley,C., Lebel,M., Pathak,S., DePinho,R.A. and Chang,S. (2005) Elevated telomere-telomere recombination in WRN-deficient, telomere dysfunctional cells promotes escape from senescence and engagement of the ALT pathway. *Genes Dev.*, **19**, 2560-2570.

17. Wyllie,F.S., Jones,C.J., Skinner,J.W., Haughton,M.F., Wallis,C., Wynford-Thomas,D., Faragher,R.G. and Kipling,D. (2000) Telomerase prevents the accelerated cell ageing of Werner syndrome fibroblasts. *Nat. Genet.*, **24**, 16-17.

18. Crabbe,L., Verdun,R.E., Haggblom,C.I. and Karlseder,J. (2004) Defective telomere lagging strand synthesis in cells lacking WRN helicase activity. *Science.*, **306**, 1951-1953.

19. Chang,S., Multani,A.S., Cabrera,N.G., Naylor,M.L., Laud,P., Lombard,D., Pathak,S., Guarente,L. and DePinho,R.A. (2004) Essential role of limiting telomeres in the pathogenesis of Werner syndrome. *Nat. Genet.*, **36**, 877-882.

20. Huang,S., Li,B., Gray,M.D., Oshima,J., Mian,I.S. and Campisi,J. (1998) The premature ageing syndrome protein, WRN, is a 3'-->5' exonuclease. *Nat. Genet.*, **20**, 114-116.

21. Bernstein,D.A. and Keck,J.L. (2003) Domain mapping of Escherichia coli RecQ defines the roles of conserved N- and C-terminal regions in the RecQ family. *Nucleic Acids Res.*, **31**, 2778-2785.

22. Lee,J.W., Kusumoto,R., Doherty,K.M., Lin,G.X., Zeng,W., Cheng,W.H., von,K.C., Brosh,R.M., Jr., Hu,J.S. and Bohr,V.A. (2005) Modulation of Werner syndrome protein function by a single mutation in the conserved RecQ domain. *J. Biol. Chem.*, **280**, 39627-39636.

23. Kitano,K., Kim,S.Y. and Hakoshima,T. (2010) Structural Basis for DNA Strand Separation by the Unconventional Winged-Helix Domain of RecQ Helicase WRN. *Structure.*, **18**, 177-187.

24. Hu,J.S., Feng,H., Zeng,W., Lin,G.X. and Xi,X.G. (2005) Solution structure of a multifunctional DNA- and protein-binding motif of human Werner syndrome protein. *Proc Natl Acad Sci U S A.*, **102,** 18379-18384.

25. Kitano,K., Yoshihara,N. and Hakoshima,T. (2007) Crystal structure of the HRDC domain of human Werner syndrome protein, WRN. *J Biol Chem.*, **282**, 2717-2728.

26. Blander,G., Kipnis,J., Leal,J.F., Yu,C.E., Schellenberg,G.D. and Oren,M. (1999) Physical and functional interaction between p53 and the Werner's syndrome protein. *J. Biol. Chem.*, **274**, 29463-29469.

27. Opresko,P.L., Mason,P.A., Podell,E.R., Lei,M., Hickson,I.D., Cech,T.R. and Bohr,V.A. (2005) POT1 stimulates RecQ helicases WRN and BLM to unwind telomeric DNA substrates. *J. Biol. Chem.*, **280**, 32069-32080.

28. Opresko,P.L., von Kobbe,C., Laine,J.P., Harrigan,J., Hickson,I.D. and Bohr,V.A. (2002) Telomere-binding protein TRF2 binds to and stimulates the Werner and Bloom syndrome helicases. *J. Biol. Chem.*, **277**, 41110-41119.

29. Opresko,P.L., Otterlei,M., Graakjaer,J., Bruheim,P., Dawut,L., Kolvraa,S., May,A., Seidman,M.M. and Bohr,V.A. (2004) The Werner syndrome helicase and exonuclease cooperate to resolve telomeric D loops in a manner regulated by TRF1 and TRF2. *Mol. Cell.*, **14**, 763-774.

30. Mohaghegh,P., Karow,J.K., Brosh,J.R., Jr., Bohr,V.A. and Hickson,I.D. (2001) The Bloom's and Werner's syndrome proteins are DNA structure-specific helicases. *Nucleic Acids Res.*, **29**, 2843-2849.

31. Shen,J.C. and Loeb,L.A. (2000) Werner syndrome exonuclease catalyzes structure-dependent degradation of DNA. *Nucleic Acids Res.*, **28**, 3260-3268.

32. Opresko,P.L., Otterlei,M., Graakjaer,J., Bruheim,P., Dawut,L., Kolvraa,S., May,A., Seidman,M.M. and Bohr,V.A. (2004) The Werner syndrome helicase and exonuclease cooperate to resolve telomeric D loops in a manner regulated by TRF1 and TRF2. *Mol. Cell*, **14**, 763-774.

33. Opresko,P.L., Laine,J.P., Brosh,R.M., Jr., Seidman,M.M. and Bohr,V.A. (2001) Coordinate action of the helicase and 3' to 5' exonuclease of Werner syndrome protein. *J. Biol. Chem.*, **276**, 44677-44687.

34. Opresko,P.L., Sowd,G. and Wang,H. (2009) The Werner syndrome helicase/exonuclease processes mobile D-loops through branch migration and degradation. *PLoS. One.*, **4**, e4825.

35. Griffith,J.D., Comeau,L., Rosenfield,S., Stansel,R.M., Bianchi,A., Moss,H. and de Lange,T. (1999) Mammalian telomeres end in a large duplex loop. *Cell*, **97**, 503-514.

36. Verdun,R.E. and Karlseder,J. (2006) The DNA damage machinery and homologous recombination pathway act consecutively to protect human telomeres. *Cell.*, **127**, 709-720.

37. Wang,R.C., Smogorzewska,A. and de Lange,T. (2004) Homologous recombination generates T-loop-sized deletions at human telomeres. *Cell.*, **119**, 355-368.

38. Stansel,R.M., de Lange,T. and Griffith,J.D. (2001) T-loop assembly in vitro involves binding of TRF2 near the 3' telomeric overhang. *EMBO J.*, **20**, 5532-5540.

39. Crabbe,L., Verdun,R.E., Haggblom,C.I. and Karlseder,J. (2004) Defective telomere lagging strand synthesis in cells lacking WRN helicase activity. *Science*, **306**, 1951-1953.

40. Poot,M., Hoehn,H., Runger,T.M. and Martin,G.M. (1992) Impaired S-phase transit of Werner syndrome cells expressed in lymphoblastoid cell lines. *Exp Cell Res.*, **202**, 267-273.

41. Huber,M.D., Duquette,M.L., Shiels,J.C. and Maizels,N. (2006) A conserved G4 DNA binding domain in RecQ family helicases. *J. Mol. Biol.*, **358**, 1071-1080.

42. Arnoult,N., Saintome,C., Ourliac-Garnier,I., Riou,J.F. and Londono-Vallejo,A. (2009) Human POT1 is required for efficient telomere C-rich strand replication in the absence of WRN. *Genes Dev.*, **23**, 2915-2924.

43. Kamath-Loeb,A.S., Loeb,L.A., Johansson,E., Burgers,P.M. and Fry,M. (2001) Interactions between the Werner syndrome helicase and DNA polymerase delta specifically facilitate copying of tetraplex and hairpin structures of the d(CGG)n trinucleotide repeat sequence. *J Biol Chem.*, **276**, 16439-16446.

44. Saintigny,Y., Makienko,K., Swanson,C., Emond,M.J. and Monnat,R.J., Jr. (2002) Homologous recombination resolution defect in werner syndrome. *Mol. Cell Biol.*, **22**, 6971-6978.

45. Wu,L. (2008) Wrestling off RAD51: a novel role for RecQ helicases. *Bioessays.*, **30**, 291-295.

46. Swanson,C., Saintigny,Y., Emond,M.J. and Monnat,R.J., Jr. (2004) The Werner syndrome protein has separable recombination and survival functions. *DNA Repair (Amst).*, **3**, 475-482.

47. Lei,M., Podell,E.R. and Cech,T.R. (2004) Structure of human POT1 bound to telomeric single-stranded DNA provides a model for chromosome end-protection. *Nat. Struct. Mol. Biol.*, **11**, 1223-1229.

48. Lei,M., Zaug,A.J., Podell,E.R. and Cech,T.R. (2005) Switching human telomerase on and off with hPOT1 protein in vitro. *J Biol Chem.*, **280**, 20449-20456.

49. Wang,F., Podell,E.R., Zaug,A.J., Yang,Y., Baciu,P., Cech,T.R. and Lei,M. (2007) The POT1-TPP1 telomere complex is a telomerase processivity factor. *Nature.*, **..**

50. Broccoli,D., Smogorzewska,A., Chong,L. and de Lange,T. (1997) Human telomeres contain two distinct Myb-related proteins, TRF1 and TRF2. *Nat. Genet.*, **17**, 231-235.

51. Fairall,L., Chapman,L., Moss,H., de Lange,T. and Rhodes,D. (2001) Structure of the TRFH dimerization domain of the human telomeric proteins TRF1 and TRF2. *Mol Cell.*, **8**, 351-361.

52. Bianchi,A., Stansel,R.M., Fairall,L., Griffith,J.D., Rhodes,D. and de Lange,T. (1999) TRF1 binds a bipartite telomeric site with extreme spatial flexibility. *EMBO J.*, **18**, 5735-5744.

53. Bianchi,A., Smith,S., Chong,L., Elias,P. and de Lange,T. (1997) TRF1 is a dimer and bends telomeric DNA. *Embo J*, **16**, 1785-1794.

54. Griffith,J., Bianchi,A. and de Lange,T. (1998) TRF1 promotes parallel pairing of telomeric tracts in vitro. *J Mol Biol.*, **278**, 79-88.

55. Hanaoka,S., Nagadoi,A. and Nishimura,Y. (2005) Comparison between TRF2 and TRF1 of their telomeric DNA-bound structures and DNA-binding activities. *Protein Sci.*, **14**, 119-130.

56. Fouche,N., Cesare,A.J., Willcox,S., Ozgur,S., Compton,S.A. and Griffith,J.D. (2006) The basic domain of TRF2 directs binding to DNA junctions irrespective of the presence of TTAGGG repeats. *J. Biol. Chem.*, **281**, 37486-37495.

57. Amiard,S., Doudeau,M., Pinte,S., Poulet,A., Lenain,C., Faivre-Moskalenko,C., Angelov,D., Hug,N., Vindigni,A., Bouvet,P. *et al.* (2007) A topological mechanism for TRF2-enhanced strand invasion. *Nat. Struct. Mol. Biol.*, **14**, 147-154.

58. Bae,N.S. and Baumann,P. (2007) A RAP1/TRF2 complex inhibits nonhomologous end-joining at human telomeric DNA ends. *Mol. Cell.*, **26**, 323-334.

59. van,S.B. and de Lange,T. (1997) Control of telomere length by the human telomeric protein TRF1. *Nature.*, **385**, 740-743.

60. Kim,S.H., Kaminker,P. and Campisi,J. (1999) TIN2, a new regulator of telomere length in human cells. *Nat Genet*, **23**, 405-412.

61. Ye,J.Z., Donigian,J.R., van,O.M., Loayza,D., Luo,Y., Krutchinsky,A.N., Chait,B.T. and de Lange,T. (2004) TIN2 binds TRF1 and TRF2 simultaneously and stabilizes the TRF2 complex on telomeres. *J Biol Chem.*, **279**, 47264-47271.

62. Maizels,N. (2006) Dynamic roles for G4 DNA in the biology of eukaryotic cells. *Nat. Struct. Mol. Biol.*, **13**, 1055-1059.

63. Balagurumoorthy,P. and Brahmachari,S.K. (1994) Structure and stability of human telomeric sequence. *J. Biol. Chem.*, **269**, 21858-21869.

64. Lee,J.Y., Okumus,B., Kim,D.S. and Ha,T. (2005) Extreme conformational diversity in human telomeric DNA. *Proc. Natl. Acad. Sci. U. S. A.*, **102**, 18938-18943.

65. Dai,J., Carver,M., Punchihewa,C., Jones,R.A. and Yang,D. (2007) Structure of the Hybrid-2 type intramolecular human telomeric G-quadruplex in K+ solution: insights into structure polymorphism of the human telomeric sequence. *Nucleic Acids Res.*, **35**, 4927-4940.

66. Schaffitzel,C., Berger,I., Postberg,J., Hanes,J., Lipps,H.J. and Pluckthun,A. (2001) In vitro generated antibodies specific for telomeric guanine-quadruplex DNA react with Stylonychia lemnae macronuclei. *Proc. Natl. Acad. Sci. U. S. A.*, **98**, 8572-8577.

67. Duquette,M.L., Handa,P., Vincent,J.A., Taylor,A.F. and Maizels,N. (2004) Intracellular transcription of G-rich DNAs induces formation of G-loops, novel structures containing G4 DNA. *Genes Dev.*, **18**, 1618-1629.

68. Parkinson,G.N., Lee,M.P. and Neidle,S. (2002) Crystal structure of parallel quadruplexes from human telomeric DNA. *Nature*, **417**, 876-880.

69. Wang,Y. and Patel,D.J. (1993) Solution structure of the human telomeric repeat d[AG3(T2AG3)3] G-tetraplex. *Structure.*, **1**, 263-282.

70. Burge,S., Parkinson,G.N., Hazel,P., Todd,A.K. and Neidle,S. (2006) Quadruplex DNA: sequence, topology and structure. *Nucleic Acids Res.*, **34**, 5402-5415.

71. Smith,F.W. and Feigon,J. (1992) Quadruplex structure of Oxytricha telomeric DNA oligonucleotides. *Nature.*, **356**, 164-168.

72. Yu,H.Q., Miyoshi,D. and Sugimoto,N. (2006) Characterization of structure and stability of long telomeric DNA G-quadruplexes. *J. Am. Chem. Soc.*, **128**, 15461-15468.

73. Marsh,T.C., Vesenka,J. and Henderson,E. (1995) A new DNA nanostructure, the G-wire, imaged by scanning probe microscopy. *Nucleic Acids Res.*, **23**, 696-700.

74. Zhang,M.L., Tong,X.J., Fu,X.H., Zhou,B.O., Wang,J., Liao,X.H., Li,Q.J., Shen,N., Ding,J. and Zhou,J.Q. (2010) Yeast telomerase subunit Est1p has guanine quadruplex-promoting activity that is required for telomere elongation. *Nat Struct Mol Biol.*, **17**, 202-209.

75. Fang,G. and Cech,T.R. (1993) The beta subunit of Oxytricha telomere-binding protein promotes G-quartet formation by telomeric DNA. *Cell.*, **74**, 875-885.

76. Eddy,J. and Maizels,N. (2008) Conserved elements with potential to form polymorphic G-quadruplex structures in the first intron of human genes. *Nucleic Acids Res.*, **36**, 1321-1333.

77. Vorlickova,M., Chladkova,J., Kejnovska,I., Fialova,M. and Kypr,J. (2005) Guanine tetraplex topology of human telomere DNA is governed by the number of (TTAGGG) repeats. *Nucleic Acids Res*, **33**, 5851-5860.

78. Yu,H.Q., Miyoshi,D. and Sugimoto,N. (2006) Characterization of structure and stability of long telomeric DNA G-quadruplexes. *J Am Chem Soc*, **128**, 15461-15468.

79. Chang,C.C., Chien,C.W., Lin,Y.H., Kang,C.C. and Chang,T.C. (2007) Investigation of spectral conversion of d(TTAGGG)4 and d(TTAGGG)13 upon potassium titration by a G-quadruplex recognizer BMVC molecule. *Nucleic Acids Res.*, **35**, 2846-2860.

80. Fouche,N., Ozgur,S., Roy,D. and Griffith,J.D. (2006) Replication fork regression in repetitive DNAs. *Nucleic Acids Res.*, **34**, 6044-6050.

81. van,S.B., Smogorzewska,A. and de Lange,T. (1998) TRF2 protects human telomeres from end-to-end fusions. *Cell.*, **92**, 401-413.

82. Fouche,N., Cesare,A.J., Willcox,S., Ozgur,S., Compton,S.A. and Griffith,J.D. (2006) The basic domain of TRF2 directs binding to DNA junctions irrespective of the presence of TTAGGG repeats. *J Biol Chem*, **281**, 37486-37495.

83. Li,B., Jog,S.P., Reddy,S. and Comai,L. (2008) WRN controls formation of extrachromosomal telomeric circles and is required for TRF2DeltaB-mediated telomere shortening. *Mol. Cell Biol.*, **28**, 1892-1904.

84. Poulet,A., Buisson,R., Faivre-Moskalenko,C., Koelblen,M., Amiard,S., Montel,F., Cuesta-Lopez,S., Bornet,O., Guerlesquin,F., Godet,T. *et al.* (2009) TRF2 promotes, remodels and protects telomeric Holliday junctions. *EMBO J.*, **28**, 641-651.

85. Kibe,T., Ono,Y., Sato,K. and Ueno,M. (2007) Fission yeast Taz1 and RPA are synergistically required to prevent rapid telomere loss. *Mol. Biol. Cell.*, **18**, 2378-2387.

86. Rog,O., Miller,K.M., Ferreira,M.G. and Cooper,J.P. (2009) Sumoylation of RecQ helicase controls the fate of dysfunctional telomeres. *Mol Cell.*, **33**, 559-569.

87. Yang,Y., Wang,H. and Erie,D.A. (2003) Quantitative characterization of biomolecular assemblies and interactions using atomic force microscopy. *Methods.*, **29**, 175-187.

88. Bustamante,C. and Rivetti,C. (1996) Visualizing protein-nucleic acid interactions on a large scale with the scanning force microscope. *Annu Rev Biophys Biomol Struct*, **25**, 395-429.

89. Hansma,H.G., Revenko,I., Kim,K. and Laney,D.E. (1996) Atomic force microscopy of long and short double-stranded, single-stranded and triple-stranded nucleic acids. *Nucleic Acids Res.*, **24**, 713-720.

90. Mikheikin,A.L., Lushnikov,A.Y. and Lyubchenko,Y.L. (2006) Effect of DNA supercoiling on the geometry of holliday junctions. *Biochemistry.*, **45**, 12998-13006.

91. Shlyakhtenko,L.S., Potaman,V.N., Sinden,R.R. and Lyubchenko,Y.L. (1998) Structure and dynamics of supercoil-stabilized DNA cruciforms. *J Mol Biol.*, **280**, 61-72.

92. Wang,H., Yang,Y., Schofield,M.J., Du,C., Fridman,Y., Lee,S.D., Larson,E.D., Drummond,J.T., Alani,E., Hsieh,P. *et al.* (2003) DNA bending and unbending by MutS govern mismatch recognition and specificity. *Proc Natl Acad Sci U S A.*, **100**, 14822-14827.

93. Wang,H., Tessmer,I., Croteau,D.L., Erie,D.A. and Van,H.B. (2008) Functional characterization and atomic force microscopy of a DNA repair protein conjugated to a quantum dot. *Nano. Lett.*, **8**, 1631-1637.

94. Neaves,K.J., Huppert,J.L., Henderson,R.M. and Edwardson,J.M. (2009) Direct visualization of G-quadruplexes in DNA using atomic force microscopy. *Nucleic Acids Res.*, **37**, 6269-75 .

95. Kan,Z.Y., Lin,Y., Wang,F., Zhuang,X.Y., Zhao,Y., Pang,D.W., Hao,Y.H. and Tan,Z. (2007) G-quadruplex formation in human telomeric (TTAGGG)4 sequence with complementary strand in close vicinity under molecularly crowded condition. *Nucleic Acids Res.*, **35**, 3646-3653.

96. Xu,Y., Ishizuka,T., Kurabayashi,K. and Komiyama,M. (2009) Consecutive formation of G-quadruplexes in human telomeric-overhang DNA: a protective capping structure for telomere ends. *Angew. Chem. Int. Ed Engl.*, **48**, 7833-7836.

97. Xu,Y. and Komiyama,M. (2007) The structural studies of human telomeric DNA using AFM. *Nucleic Acids Symp. Ser. (Oxf).*, 241-242.

98. Tang,J., Kan,Z.Y., Yao,Y., Wang,Q., Hao,Y.H. and Tan,Z. (2008) G-quadruplex preferentially forms at the very 3' end of vertebrate telomeric DNA. *Nucleic Acids Res.*, **36**, 1200-1208.

99. Cornish,P.V. and Ha,T. (2007) A survey of single-molecule techniques in chemical biology. *ACS Chem. Biol.*, **2**, 53-61.

100. Burge,S., Parkinson,G.N., Hazel,P., Todd,A.K. and Neidle,S. (2006) Quadruplex DNA: sequence, topology and structure. *Nucleic Acids Res*, **34**, 5402-5415.

101. Li,J., Correia,J.J., Wang,L., Trent,J.O. and Chaires,J.B. (2005) Not so crystal clear: the structure of the human telomere G-quadruplex in solution differs from that present in a crystal. *Nucleic Acids Res.*, **33**, 4649-4659.

102. Zaug,A.J., Podell,E.R. and Cech,T.R. (2005) Human POT1 disrupts telomeric G-quadruplexes allowing telomerase extension in vitro. *Proc. Natl. Acad. Sci. U. S. A.*, **102**, 10864-10869.

103. Zaug,A.J., Podell,E.R. and Cech,T.R. (2005) Human POT1 disrupts telomeric G-quadruplexes allowing telomerase extension in vitro. *Proc Natl Acad Sci U S A*, **102**, 10864-10869.

104. Dai,J., Carver,M., Punchihewa,C., Jones,R.A. and Yang,D. (2007) Structure of the Hybrid-2 type intramolecular human telomeric G-quadruplex in K+ solution: insights into structure polymorphism of the human telomeric sequence. *Nucleic Acids Res.*, **35**, 4927-4940.

105. Petraccone,L., Trent,J.O. and Chaires,J.B. (2008) The tail of the telomere. *J Am Chem Soc.*, **130**, 16530-16532.

106. Lyubchenko,Y.L. and Shlyakhtenko,L.S. (1997) Visualization of supercoiled DNA with atomic force microscopy in situ. *Proc. Natl. Acad. Sci. U. S. A.*, **94**, 496-501.

107. Hansma,H.G., Revenko,I., Kim,K. and Laney,D.E. (1996) Atomic force microscopy of long and short double-stranded, single-stranded and triple-stranded nucleic acids. *Nucleic Acids Res.*, **24**, 713-720.

108. Hockemeyer,D., Sfeir,A.J., Shay,J.W., Wright,W.E. and de Lange,T. (2005) POT1 protects telomeres from a transient DNA damage response and determines how human chromosomes end. *EMBO J.*, **24**, 2667-2678.

109. White,J.H. and Bauer,W.R. (2004) Finite-element analysis of the displacement of closed DNA loops under torsional stress. *Philos. Transact. A Math. Phys. Eng Sci.*, **362**, 1335-1353.

110. Wang,H., Yang,Y. and Erie,D.A. (2007) Characterization of protein-protein interactions using atomic force microscopy. In Schuck,P. (ed.), *Protein Interactions Biophysical approaches for the study of complex reversible systems*. Springer Science+Business Media, LLC, pp. 39-78.

111. Wang,H., Tessmer,I., Croteau,D.L., Erie,D.A. and Van,H.B. (2008) Functional characterization and atomic force microscopy of a DNA repair protein conjugated to a quantum dot. *Nano. Lett.*, **8**, 1631-1637.

112. Zhang,X.Y., Cao,E.H., Zhang,Y., Chou,C. and Bai,C. (2003) K+ and Na+-induced self-assembly of telomeric oligonucleotide d(TTAGGG)n. *J Biomol Struct Dyn*, **20**, 693-702.

113. Yoshimura,S.H., Maruyama,H., Ishikawa,F., Ohki,R. and Takeyasu,K. (2004) Molecular mechanisms of DNA end-loop formation by TRF2. *Genes Cells*, **9**, 205-218.

114. Xu,Y., Ishizuka,T., Kurabayashi,K. and Komiyama,M. (2009) Consecutive formation of G-quadruplexes in human telomeric-overhang DNA: a protective capping structure for telomere ends. *Angew Chem Int Ed Engl*, **48**, 7833-7836.

115. Zhang,X.Y., Cao,E.H., Zhang,Y., Chou,C. and Bai,C. (2003) K+ and Na+-induced self-assembly of telomeric oligonucleotide d(TTAGGG)n. *J. Biomol. Struct. Dyn.*, **20**, 693-702.

116. Yang,Y., Wang,H. and Erie,D.A. (2003) Quantitative characterization of biomolecular assemblies and interactions using atomic force microscopy. *Methods*, **29**, 175-187.

117. Ambrus,A., Chen,D., Dai,J., Bialis,T., Jones,R.A. and Yang,D. (2006) Human telomeric sequence forms a hybrid-type intramolecular G-quadruplex structure with mixed parallel/antiparallel strands in potassium solution. *Nucleic Acids Res.*, **34**, 2723-2735.

118. Renciuk,D., Kejnovska,I., Skolakova,P., Bednarova,K., Motlova,J. and Vorlickova,M. (2009) Arrangements of human telomere DNA quadruplex in physiologically relevant K+ solutions. *Nucleic Acids Res.*, **37**, 6625-6634.

119. Petraccone,L., Garbett,N., Chaires,J.B. and Trent,J.O. (2010) An integrated molecular dynamics and experimental study of higher order human telomeric quadruplexes. *Biopolymers.*.

120. Marsh,T.C., Vesenka,J. and Henderson,E. (1995) A new DNA nanostructure, the G-wire, imaged by scanning probe microscopy. *Nucleic Acids Res*, **23**, 696-700.

121. Lei,M., Podell,E.R. and Cech,T.R. (2004) Structure of human POT1 bound to telomeric single-stranded DNA provides a model for chromosome end-protection. *Nat Struct Mol Biol*, **11**, 1223-1229.

122. Ratcliff,G.C. and Erie,D.A. (2001) A novel single-molecule study to determine protein--protein association constants. *J Am Chem Soc*, **123**, 5632-5635.

123. Lee,J.Y., Okumus,B., Kim,D.S. and Ha,T. (2005) Extreme conformational diversity in human telomeric DNA. *Proc Natl Acad Sci U S A*, **102**, 18938-18943.

124. Torigoe,H. and Furukawa,A. (2007) Tetraplex structure of fission yeast telomeric DNA and unfolding of the tetraplex on the interaction with telomeric DNA binding protein Pot1. *J Biochem (Tokyo)*, **141**, 57-68.

125. Trujillo,K.M., Bunch,J.T. and Baumann,P. (2005) Extended DNA binding site in Pot1 broadens sequence specificity to allow recognition of heterogeneous fission yeast telomeres. *J Biol Chem*, **280**, 9119-9128.

126. Lei,M., Zaug,A.J., Podell,E.R. and Cech,T.R. (2005) Switching human telomerase on and off with hPOT1 protein in vitro. *J Biol Chem*, **280**, 20449-20456.

127. Verdun,R.E., Crabbe,L., Haggblom,C. and Karlseder,J. (2005) Functional human telomeres are recognized as DNA damage in G2 of the cell cycle. *Mol Cell*, **20**, 551-561.

128. Wu,Y., Shin-ya,K. and Brosh,R.M., Jr. (2008) FANCJ helicase defective in Fanconia anemia and breast cancer unwinds G-quadruplex DNA to defend genomic stability. *Mol Cell Biol*, **28**, 4116-4128.

129. Bhattacharyya,S., Sandy,A. and Groden,J. (2010) Unwinding protein complexes in ALTernative telomere maintenance. *J Cell Biochem*, **109**, 7-15.

130. London,T.B., Barber,L.J., Mosedale,G., Kelly,G.P., Balasubramanian,S., Hickson,I.D., Boulton,S.J. and Hiom,K. (2008) FANCJ is a structure-specific DNA helicase associated with the maintenance of genomic G/C tracts. *J Biol Chem*, **283**, 36132-36139.

131. Rizzo,A., Salvati,E., Porru,M., D'Angelo,C., Stevens,M.F., D'Incalci,M., Leonetti,C., Gilson,E., Zupi,G. and Biroccio,A. (2009) Stabilization of quadruplex DNA perturbs telomere replication leading to the activation of an ATR-dependent ATM signaling pathway. *Nucleic Acids Res*, **37**, 5353-5364.

132. Sowd,G., Lei,M. and Opresko,P.L. (2008) Mechanism and substrate specificity of telomeric protein POT1 stimulation of the Werner syndrome helicase. *Nucleic Acids Res*, **36**, 4242-4256.

133. Chu,W.K. and Hickson,I.D. (2009) RecQ helicases: multifunctional genome caretakers. *Nat. Rev. Cancer*, **9**, 644-654.

134. Sidorova,J.M., Li,N., Folch,A. and Monnat,R.J., Jr. (2008) The RecQ helicase WRN is required for normal replication fork progression after DNA damage or replication fork arrest. *Cell Cycle.*, **7**, 796-807.

135. Crabbe,L., Jauch,A., Naeger,C.M., Holtgreve-Grez,H. and Karlseder,J. (2007) Telomere dysfunction as a cause of genomic instability in Werner syndrome. *Proc. Natl. Acad. Sci. U.S.A.*, **104**, 2205-2210.

136. Constantinou,A., Tarsounas,M., Karow,J.K., Brosh,R.M., Bohr,V.A., Hickson,I.D. and West,S.C. (2000) Werner's syndrome protein (WRN) migrates Holliday junctions and co-localizes with RPA upon replication arrest. *EMBO Rep.*, **1**, 80-84.

137. Makarov,V.L., Hirose,Y. and Langmore,J.P. (1997) Long G tails at both ends of human chromosomes suggest a C strand degradation mechanism for telomere shortening. *Cell*, **88**, 657-666.

138. Karlseder,J., Hoke,K., Mirzoeva,O.K., Bakkenist,C., Kastan,M.B., Petrini,J.H. and de Lange,T. (2004) The telomeric protein TRF2 binds the ATM kinase and can inhibit the ATM-dependent DNA damage response. *PLoS. Biol.*, **2**, E240.

139. Denchi,E.L. and de Lange,T. (2007) Protection of telomeres through independent control of ATM and ATR by TRF2 and POT1. *Nature*, **448**, 1068-1071.

140. Cesare,A.J. and Griffith,J.D. (2004) Telomeric DNA in ALT cells is characterized by free telomeric circles and heterogeneous t-loops. *Mol. Cell Biol.*, **24**, 9948-9957.

141. Shen,J.C. and Loeb,L.A. (2000) Werner syndrome exonuclease catalyzes structure-dependent degradation of DNA. *Nucleic Acids Res.*, **28**, 3260-3268.

142. Compton,S.A., Tolun,G., Kamath-Loeb,A.S., Loeb,L.A. and Griffith,J.D. (2008) The Werner syndrome protein binds replication fork and holliday junction DNAs as an oligomer. *J. Biol. Chem.*, **283**, 24478-24483.

143. Machwe,A., Xiao,L. and Orren,D.K. (2004) TRF2 recruits the Werner syndrome (WRN) exonuclease for processing of telomeric DNA. *Oncogene*, **23**, 149-156.

144. Sowd,G., Lei,M. and Opresko,P.L. (2008) Mechanism and substrate specificity of telomeric protein POT1 stimulation of the Werner syndrome helicase. *Nucleic Acids Res.*, **36**, 4242-4256.

145. Zhitkovich,A. and Costa,M. (1992) A simple, sensitive assay to detect DNA-protein crosslinks in intact cells and in vivo. *Carcinogenesis.*, **13**, 1485-1489.

146. Machwe,A., Xiao,L., Groden,J. and Orren,D.K. (2006) The Werner and Bloom syndrome proteins catalyze regression of a model replication fork. *Biochemistry.*, **45**, 13939-13946.

147. Brosh,R.M., Jr., Driscoll,H.C., Dianov,G.L. and Sommers,J.A. (2002) Biochemical characterization of the WRN-FEN-1 functional interaction. *Biochemistry.*, **41**, 12204-12216.

148. Li,B., Oestreich,S. and de Lange,T. (2000) Identification of human Rap1: implications for telomere evolution. *Cell*, **101**, 471-483.

149. Singleton,M.R., Scaife,S. and Wigley,D.B. (2001) Structural analysis of DNA replication fork reversal by RecG. *Cell*, **107**, 79-89.

150. Pyle,A.M. (2008) Translocation and unwinding mechanisms of RNA and DNA helicases. *Annu. Rev. Biophys.*, **37:317-36.**, 317-336.

151. Lloyd,R.G. and Sharples,G.J. (1993) Processing of recombination intermediates by the RecG and RuvAB proteins of Escherichia coli. *Nucleic Acids Res.*, **21**, 1719-1725.

152. Laud,P.R., Multani,A.S., Bailey,S.M., Wu,L., Ma,J., Kingsley,C., Lebel,M., Pathak,S., DePinho,R.A. and Chang,S. (2005) Elevated telomere-telomere recombination in WRN-deficient, telomere dysfunctional cells promotes escape from senescence and engagement of the ALT pathway. *Genes Dev.*, **19**, 2560-2570.

153. Fairall,L., Chapman,L., Moss,H., de Lange,T. and Rhodes,D. (2001) Structure of the TRFH dimerization domain of the human telomeric proteins TRF1 and TRF2. *Mol. Cell.*, **8**, 351-361.

154. Yang,Q., Zhang,R., Wang,X.W., Spillare,E.A., Linke,S.P., Subramanian,D., Griffith,J.D., Li,J.L., Hickson,I.D., Shen,J.C. *et al.* (2002) The processing of Holliday junctions by BLM and WRN helicases is regulated by p53. *J. Biol. Chem.*, **277**, 31980-31987.

155. Khuu,P.A., Voth,A.R., Hays,F.A. and Ho,P.S. (2006) The stacked-X DNA Holliday junction and protein recognition. *J. Mol. Recognit.*, **19**, 234-242.

156. Karymov,M., Daniel,D., Sankey,O.F. and Lyubchenko,Y.L. (2005) Holliday junction dynamics and branch migration: single-molecule analysis. *Proc. Natl. Acad. Sci. U.S.A.*, **102**, 8186-8191.

157. McKinney,S.A., Freeman,A.D., Lilley,D.M. and Ha,T. (2005) Observing spontaneous branch migration of Holliday junctions one step at a time. *Proc. Natl. Acad. Sci. U.S.A.*, **102**, 5715-5720.

158. Parsons,C.A., Stasiak,A., Bennett,R.J. and West,S.C. (1995) Structure of a multisubunit complex that promotes DNA branch migration. *Nature*, **374**, 375-378.

159. Sarthy,J., Bae,N.S., Scrafford,J. and Baumann,P. (2009) Human RAP1 inhibits non-homologous end joining at telomeres. *EMBO J.*, **28**, 3390-3399.

160. Li,B., Oestreich,S. and de Lange,T. (2000) Identification of human Rap1: implications for telomere evolution. *Cell*, **101**, 471-483.

161. de Lange,T. (2009) How telomeres solve the end-protection problem. *Science*, **326**, 948-952.

162. Rog,O., Miller,K.M., Ferreira,M.G. and Cooper,J.P. (2009) Sumoylation of RecQ helicase controls the fate of dysfunctional telomeres. *Mol. Cell.*, **33**, 559-569.

163. Kibe,T., Ono,Y., Sato,K. and Ueno,M. (2007) Fission yeast Taz1 and RPA are synergistically required to prevent rapid telomere loss. *Mol. Biol. Cell.*, **18**, 2378-2387.

164. Tang,J., Kan,Z.Y., Yao,Y., Wang,Q., Hao,Y.H. and Tan,Z. (2008) G-quadruplex preferentially forms at the very 3' end of vertebrate telomeric DNA. *Nucleic Acids Res*, **36**, 1200-1208.

165. Barrientos,K.S., Kendellen,M.F., Freibaum,B.D., Armbruster,B.N., Etheridge,K.T. and Counter,C.M. (2008) Distinct functions of POT1 at telomeres. *Mol Cell Biol.*, **28**, 5251-5264.

166. Ackerson,C.J., Jadzinsky,P.D., Sexton,J.Z., Bushnell,D.A. and Kornberg,R.D. (2010) Synthesis and Bioconjugation of 2 and 3 nm-Diameter Gold Nanoparticles. *Bioconjug. Chem.*, **21**, 214-218.

167. Mergny,J.L. and Maurizot,J.C. (2001) Fluorescence resonance energy transfer as a probe for G-quartet formation by a telomeric repeat. *Chembiochem.*, **2**, 124-132.